AD 732795

Technical Report: NAVTRADEVCEN 69-C-0322-2

TRAINING EFFECTIVENESS EVALUATION OF NAVAL TRAINING DEVICES PART II: A STUDY OF DEVICE 2F66A (S-2E TRAINER) EFFECTIVENESS

David Meister
D.J. Sullivan
E.A. Thompson
D.L. Finley

The Bunker-Ramo Corporation
Electronic Systems Division
31717 La Tienda Drive
Westlake Village, California 91361

Contract N61339-69-C-0322
NAVTRADEVCEN Task No. 8264

July 1971

DOD DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

NAVAL TRAINING DEVICE CENTER

ORLANDO, FLORIDA

107

Technical Report: NAVTRADEVCEN 69-C-0322-2

TRAINING EFFECTIVENESS EVALUATION OF NAVAL TRAINING DEVICES

ABSTRACT

A study was performed of the effectiveness of Device 2F66A (S-2E aircraft) in training crew members to perform air antisubmarine warfare missions.

Although each conclusion must be qualified by consideration of the data sample and detailed results described in the report, it was possible to arrive at the following overall conclusions:

a. The trainer was found to be effective in training all crew members. The data suggests it was more effective for No. 4 operators and less effective for No. 3 operators and TACCOs.

b. The trainer appeared to be more effective for beginning students than for operational and reserve personnel.

c. Factors which instructors used to vary the difficulty of training sessions did not appear to have a major effect on overall trainee performance.

d. The effect of team training sessions on student performance was consistent only during the early part of the first team session.

e. Instructors considered the trainer effective in training. However, equipment malfunctions and insufficient fidelity have reduced confidence and lowered the overall acceptability of the device as an integral part of air ASW training.

Recommendations for improved use of existing capabilities are discussed. To summarize, it is suggested that trainer effectiveness for operational objectives could be increased by: (1) more systematic variation of the nature of the training session, (2) increased trainer usage, and (3) more systematic utilization of the trainer in both individual position and team training modes with suggested changes in mode emphasis, and (4) improved trainer maintenance, enhanced fidelity, and increased trainer hardware capability. Further, specific recommendations were made for improvement of personnel motivation and training of the individual team member. Areas requiring further investigation, both for determination of the most effective training program for operational needs and for evaluation of certain hypotheses currently utilized for training program development are delineated.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) The Bunker-Ramo Corporation 31717 La Tienda Drive Westlake Village, California 91306	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE
Training Effectiveness Evaluation of Naval Training Devices, Part II: A Study of Device 2F66A(S-2E Trainer) Effectiveness

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)
David Meister D. L. Finley
D. J. Sullivan
E. A. Thompson

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
July 1971	112	6

8a. CONTRACT OR GRANT NO. N61339-69-C-0322	9a. ORIGINATOR'S REPORT NUMBER(S) J93-1U1
b. PROJECT NO. Project No. 8264	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	NAVTRADEVCEN 69-C-0322-2

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY U.S. Naval Training Device Center Orlando, Florida

13. ABSTRACT

This study examined the effectiveness of Device 2F66A (S-2E aircraft) in training crew members to perform air antisubmarine warfare missions. Data collection was made through daily on-site observations, and four types of data were collected.

Results generally indicated that the trainer was effective with all crew members, that beginning students benefitted more than operational and reserve personnel, and that important improvements in training for all personnel could be effected through improved, systematic utilization of existing training devices.

Recommendations for improved use of capabilities include: more systematic variations of the training session, increased trainer usage, and more systematic utilization of the trainer in both individual positions and team training modes.

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Training Effectiveness						
Training Devices						
AASW Training						
S-2E Training						

Unclassified
Security Classification

FOREWORD

The problem of evaluating the effectiveness of a training system in aiding the attainment of proficiency in a complex task is one which requires careful analyses of the many variables that affect the operational task. Once these variables have been identified and their effects on performance analyzed, objective measurement schemes can be devised to record pertinent data, the analysis of which will indicate the amount of learning that has taken place in the training situation. Each training effectiveness evaluation performed in the field is somewhat unique and generally provides insight into new methods for obtaining valid performance measures. Such evaluations also tell us something about the usefulness of types of performance criteria. As noted in the present study, not all measures taken were indicators of learning. The experience derived from such studies should eventually produce a rationale for choosing measures that are meaningful and rejecting ones that are not. It is expected that by analyzing the results of a series of field evaluations, measurement techniques may be made more sensitive, and refined to the point where they can be applied more efficiently and made to produce more reliable results.

One interesting finding of the present study was the difference in performance found between classes of trainees (i.e., students, operational, and reserve personnel). Another significant finding was afforded by comparison of the results from the same type of trainers located at two different facilities (San Diego and Key West). As indicated in the body of this report, several factors were probably responsible for the differences found. The important point to be emphasized, however, is the influence of variables outside the direct control of the training device itself. This illustrates the importance of studying the training system as a whole, including the curriculum, training personnel, trainee population, and training environment. As shown by this study, the same training device may produce differences in training as a result of any one, or an interaction, of these outside forces.

Joseph A. Puig

JOSEPH A. PUIG
Project Psychologist
Naval Training Device Center

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont'd)

SECTION I

DESCRIPTION OF THE STUDY

PURPOSE OF THE STUDY

This report documents the results of one of a series of studies being conducted for the Naval Training Devices Center under the project title "Training Effectiveness Evaluation for Naval Training Devices" (Project 8264).

The question to which this investigation was addressed ("Is this training device effective?") must be translated into more specific study objectives. These are discussed in the paragraphs to follow.

DO STUDENTS TRAINED ON DEVICE 2F66A SHOW EVIDENCE OF LEARNING? Learning on a training device can be measured in a number of ways. The most desirable method is to demonstrate transfer of training to the operational situation (Jeantheau's (1970), level IV analysis). To do so, however, requires the opportunity to measure performance of the student in that operational setting. In the present investigation the opportunity to obtain transfer measures was lacking. However, the following indices of learning, although less desirable than the transfer of training measure, do provide useful information regarding training device effectiveness:

a. A substantial improvement in performance in the course of training over performance in the first training session

b. A tendency on the part of the learning curve to show progressive improvement

c. Student performance reaching or exceeding the mean performance of operational personnel at the conclusion of the former's training.

Performance measurement data must therefore be collected on a sequence of training sessions for at least student and operational personnel such that the effects of continued trainer use can be described and evaluated.

HOW DOES LEARNING COMPARE AMONG BEGINNING STUDENTS, OPERATIONAL PERSONNEL, AND RESERVIST PERSONNEL? Because the subject population for this study consisted of the three different samples noted above, it was necessary to determine the differential effect of the trainer on the performance of each sample. Conceivably the trainer could be more effective for one type of student than another.

HOW IS LEARNING ON THE TRAINING DEVICE INFLUENCED BY MANIPULATION OF TRAINING INPUTS? The inputs that we refer to are those which make up the problems presented to trainees. Examples of such inputs are: the number of targets presented, target speed, or simulated equipment failures.

The intended effect of such inputs is to vary the level of difficulty of the training sessions. Training is generally considered more effective when the level of difficulty of each successive training session is increased progressively. It is therefore necessary to examine the manner in which these difficulty inputs are included in the training situation. Conceivably substantial improvements in trainer effectiveness can be achieved by adjusting input difficulty to student requirements.

HOW IS LEARNING PERFORMANCE ON THE TRAINER INFLUENCED BY THE INTRODUCTION OF TEAM TRAINING AND BY THE SEQUENCING OF SIMULATOR AND FLIGHT TRAINING SESSIONS? Although the device under evaluation was designed to be a Weapon System Trainer (WST) to train the entire crew in coordinated mission operations, it is also used to train crew members in their individual duties. It is therefore reasonable to ask whether the introduction of team training sessions after the student has been trained in his individual duties and tasks has any effect on the student's performance.

Training, using the 2F66A, is performed either concurrently with or immediately preceding flight training. Since it is conceivable that flight training may serve to enhance or inhibit ground training, it is desirable to explore the relationship between the two in terms of the learning achieved in the trainer.

HOW IS THE TRAINING DEVICE EVALUATED BY ITS INSTRUCTORS? It seems reasonable to assume that if the instructors who make use of a training device consider it effective in training, its utility in training will be increased. Since the personnel who instruct on a device probably have the most intimate knowledge of its strong and weak points, it seems desirable to secure their evaluations of its effectiveness. Knowledge of these device characteristics which instructors view as most and least desirable could also be fed back into the design of new training devices.

TRAINING DEVICE DESCRIPTION AND OPERATIONAL SETTING

Device 2F66A is designed to function as a combined air antisubmarine warfare (AASW) tactics and operational flight trainer (OFT), providing training in the wide range of tactical and operational capabilities of the Grumman S-2E aircraft. The S-2E aircraft is a four-place, twin engine, high wing monoplane designed for use in antisubmarine warfare; the functions of the aircraft are to search for, detect, track, localize, and destroy submarines under all weather flight conditions. In the context of this study we are concerned only with the weapon system aspects of the trainer, i.e., that training provided in AASW tactics.

THE AASW MISSION. The AASW mission in gross terms is an extremely variable one. To a great extent its nature is determined not only by the characteristics and capabilities of the aircraft performing the mission, but also by the nature of the tactical situation, e.g., target maneuvers. This mission can be defined in terms of two categories, search and attack:

Search is the systematic investigation of a particular area for the purpose of locating, or confirming the absence of, an object suspected of being in that area.

Attack is the logical prosecution of the search phase using the weapons available in accordance with developed tactics, criteria, and procedures.

The majority of operational S-2E mission time is devoted to the search function of its mission. Within both functions it is possible to define a logical sequence of activities whose nature and order are defined by the capabilities and constraints of the sensor and processing systems available. An analysis of the mission, functions, and tasks assigned to the S-2E identified the following phases which are common to all missions. These are:

a. LOFAR or Barrier search

b. CODAR search

c. JULIE localization

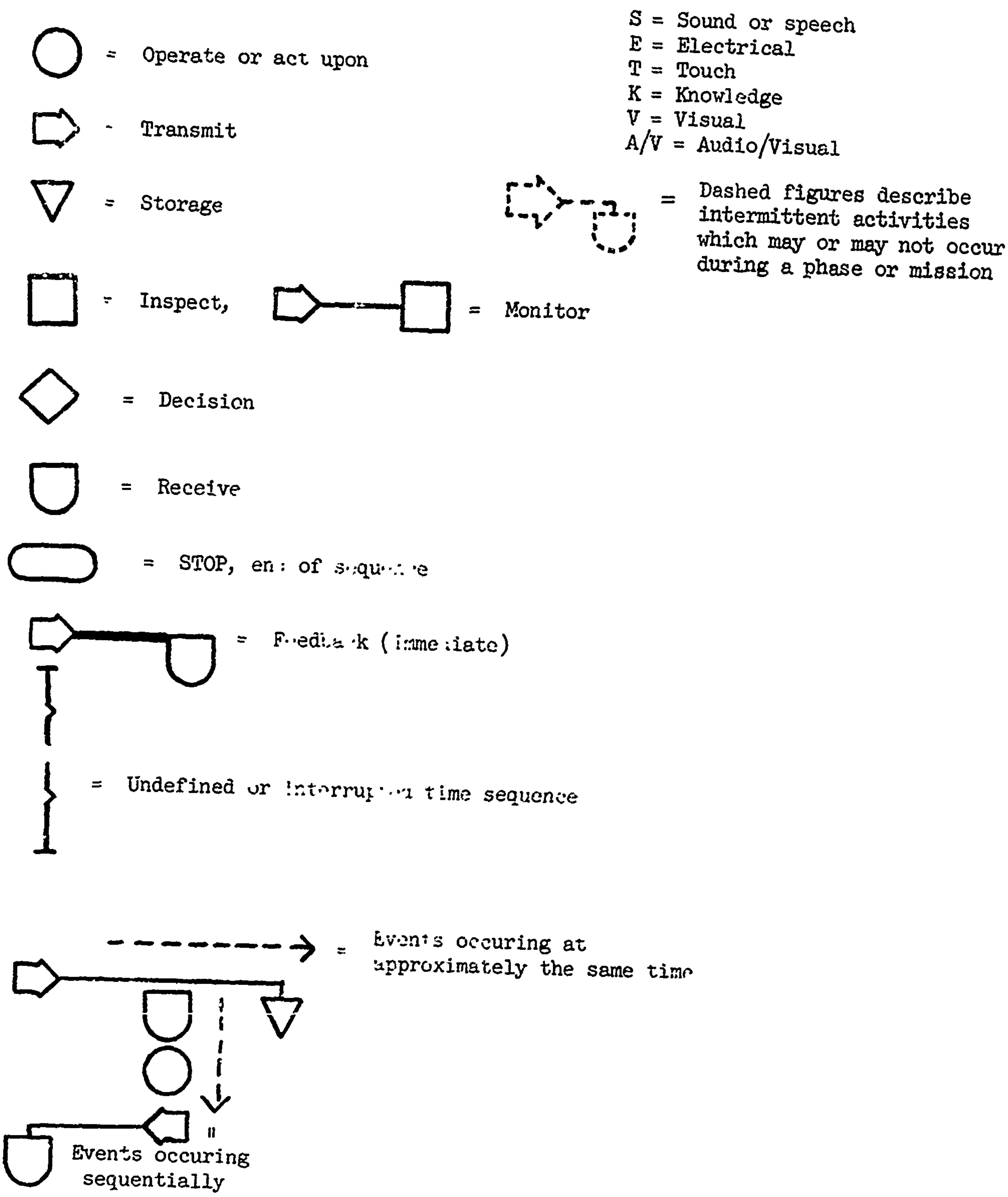
d. MAD search

e. Attack.

All complete operational missions, regardless of individual variations, must contain the listed phases conducted in the above sequence. However, if contact with the target is lost, return to an earlier phase and reiteration of the sequence may occur. In addition, any one mission may not contain all of the phases. For example, it is possible to conduct either a LOFAR or Barrier search, to go directly into JULIE, and then attack, eliminating the use of CODAR and MAD phases.

Training missions in device 2F66A simulated operational missions; emphasis on particular phases was induced by training inputs at the instructor's discretion (pp 11-14). As part of the analysis leading up to the identification of these standard mission phases, an Operational Sequence Diagram (OSD) (Figure 1) was developed which identifies the relationship of one mission segment to another as well as identifying the functions of individual crew members in each mission phase. The purpose of the OSD was to provide a basis for defining measures of individual and team performance relevant to operational S-2E mission objectives.

OSD LEGEND



LEGEND FOR FIGURE 1.

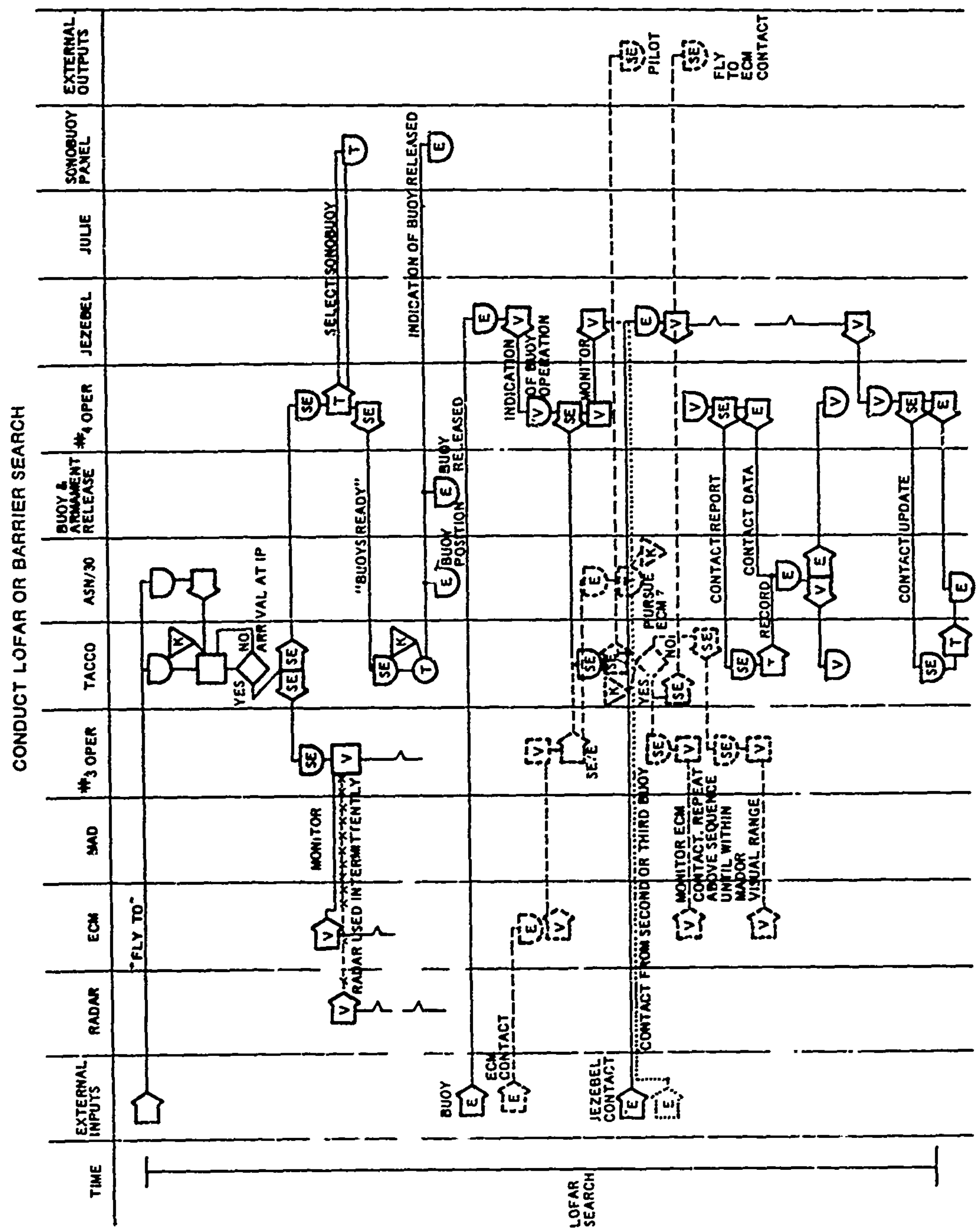


Figure 1. Operational Sequence Diagram of AASW Mission (Part 1 of 4)

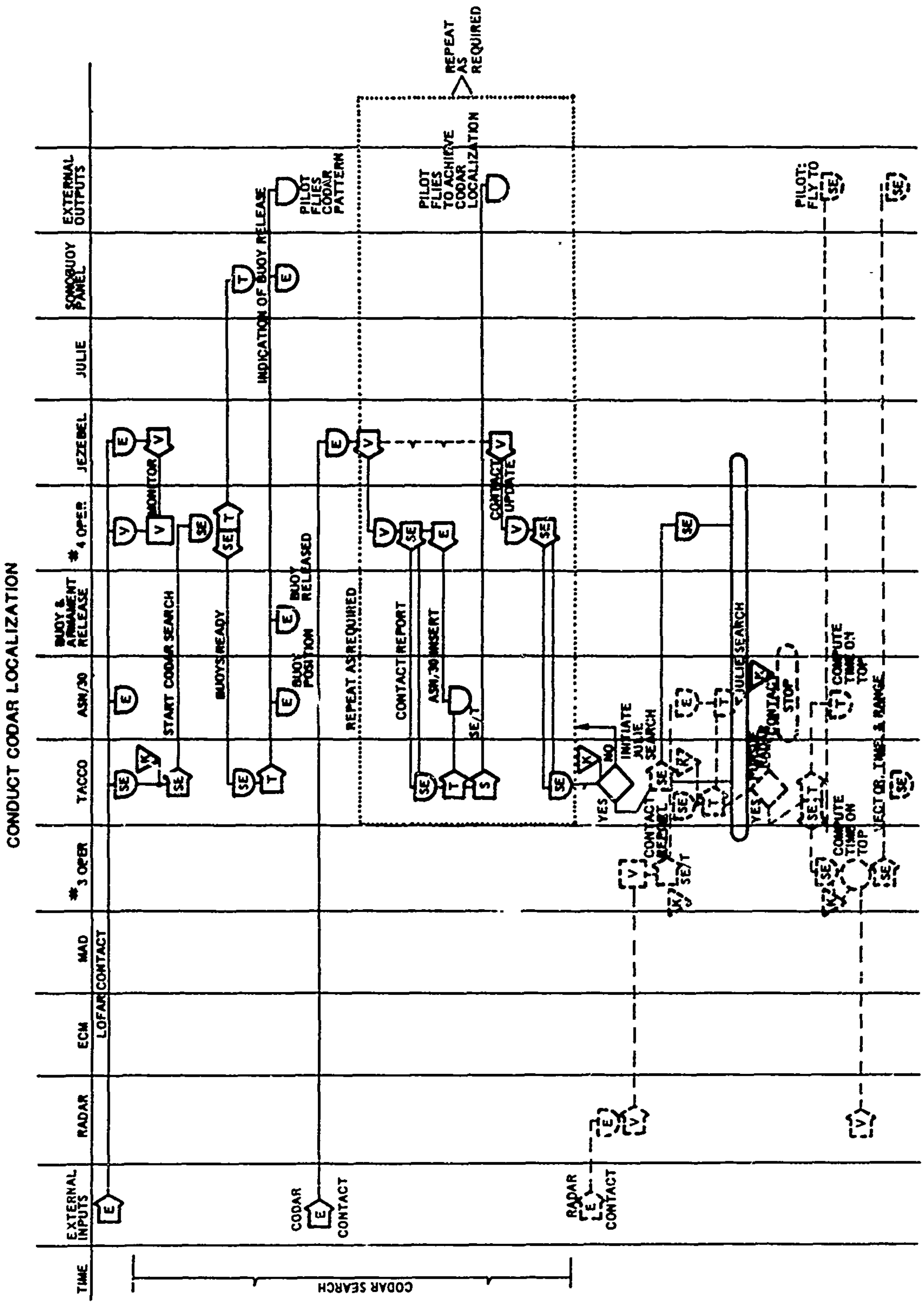


Figure 1. Operational Sequence Diagram of AASW Mission (Part 2 of 4)

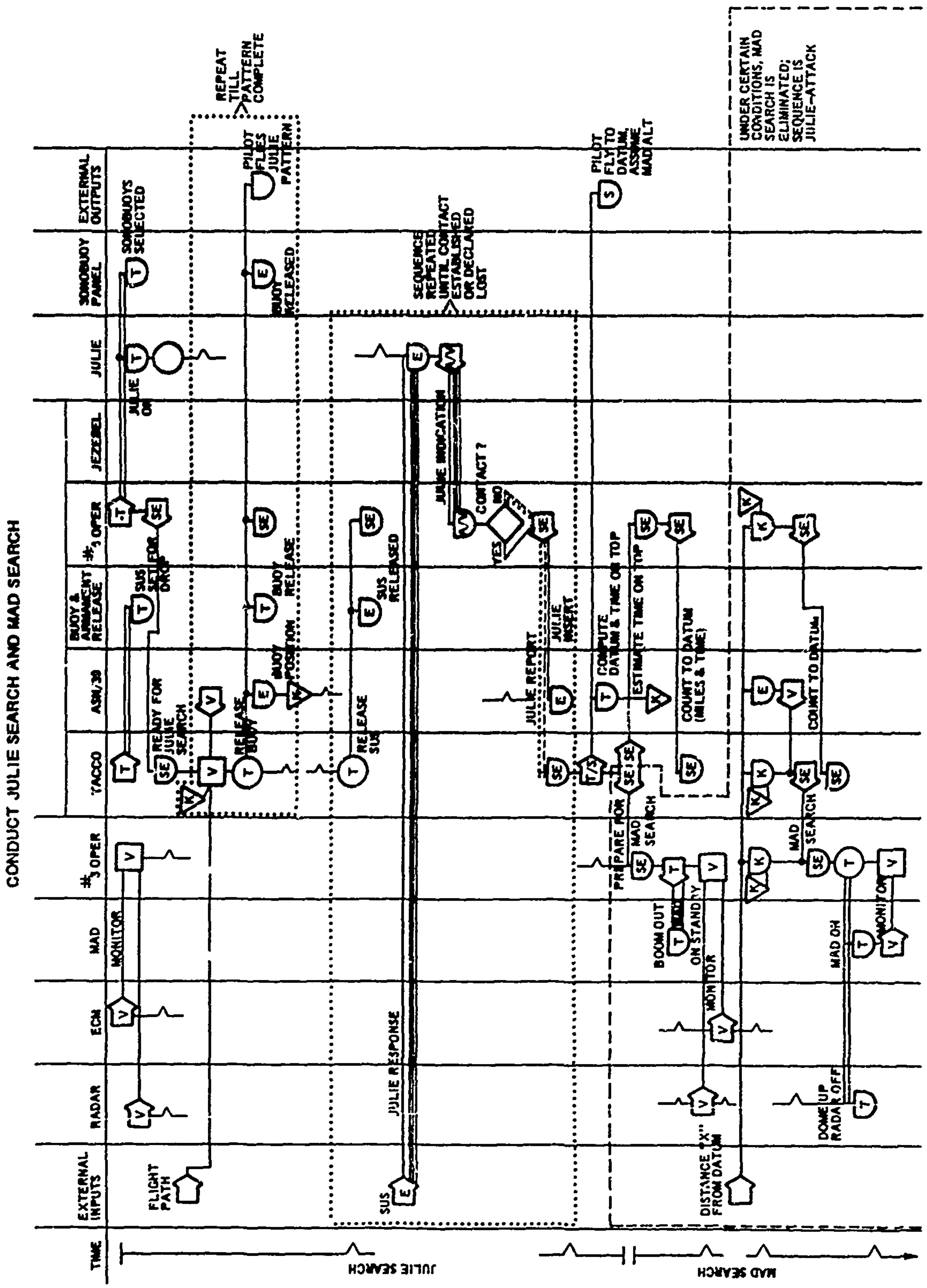


Figure 1. Operational Sequence Diagram of AASW Mission (Part 3 of 4)

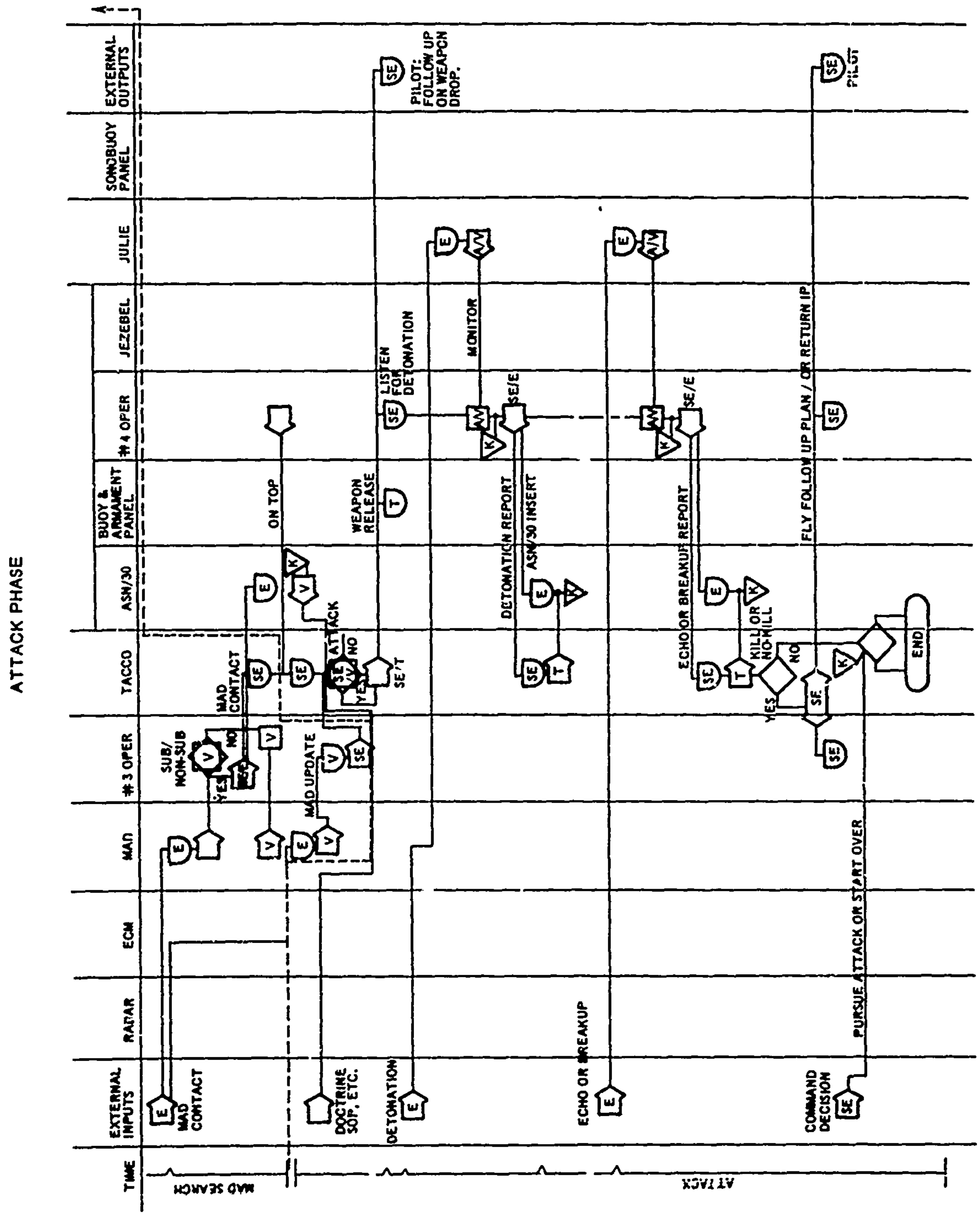


Figure 1. Operational Sequence Diagram of AASW Mission (Part 4 of 4)

S-2E CREW POSITIONS AND TASKS. There are four positions in the S-2E aircraft. The functional roles of each of the four individuals who make up the flight crew are listed below:

Pilot. This individual is the aircraft commander and he is normally responsible for operation of the aircraft controls. His duties include, but are not limited to:

a. Mission planning and briefing

b. Pre-flight inspection

c. Performance of starting procedures, system and control checks, taxiing, engine run-up, take-off, landing and ground secure

d. Instructing the tactical coordinator of the mission relative to the assistance he desires in maintaining control of the aircraft.

Although the pilot's flight activities impact upon the performance of the mission (in the sense that, for example, flying the wrong heading or the wrong altitude would result in failure to detect or track the target), they are not integral to the AASW mission and hence were not considered in the study. The pilot was considered, however, in analyzing crew communications.

TACCO. The term describes the individual who serves as tactical coordinator of the mission. The TACCO's duties include:

*a. Performing all communications functions

b. Maintaining an up-to-date navigational plot as required by the mission

*c. Performing the duties of the tactical coordinator, maintaining an up-to-date tactical plot and such logs as may be required by the mission and tactics involved

d. Assisting the pilot by monitoring the instruments.

No. 3 Operator. This term is used to identify the person in the left seat of the crew compartment. His duties involve operation of the nonacoustic equipment aboard the S-2E and include the following.

---

*Asterisks identify the AASW operations which were considered in the evaluation of the trainer.

*a. Operating and reporting the status of the Magnetic Anomaly Detector (MAD), electronic countermeasures (ECM), Radar and Navigation Computer.

b. Troubleshooting the above listed gear.

c. Maintaining a visual lookout as permitted by his other duties.

No. 4 Operator. This term identifies the person in the right seat of the crew compartment. His duties involve operation of the acoustic equipment aboard the S-2E and include:

*a. Operating and reporting the status of the JULIE and JEZEBEL recorders, sonobuoy receivers, and ordnance panels

b. Troubleshooting the equipment listed above

c. Maintaining a visual lookout as permitted by his other duties.

No description of the equipment referred to above (e.g., MAD, ECM) has been provided, except as necessary for an understanding of the performance measures recorded as part of the study.

TRAINING DEVICE DESCRIPTION. The S-2E Weapon System Trainer, Device 2F66A, was manufactured by the Link Group, Systems Division of General Precision Inc. The device is designed to be an accurate reproduction of the pilot, TACCO, and operator positions of the Grumman S-2E aircraft and to provide a complete and realistic environment for trainees to "fly" any type of mission which can be accomplished by the actual aircraft. For further information on the device see NAVSO P-2853-R1, February 1967.

Device 2F66A consists of two 40-foot semi-trailers. One trailer simulates the flight crew and operators' stations, the instructors' console and associated control panels; the second trailer contains all the computing equipment necessary to simulate the flight and tactical equipment and instruments present in the other trailer.

Flight Trainees' Compartment. The consoles at the flight trainees' stations are exact physical replicas of those in the actual aircraft cockpit. With a few exceptions, the operational components and displays are functional replicas of the operational equipment.

*Asterisks identify the AASW operations which were considered in the evaluation of the trainer.

<u>Tactics Trainees' Compartment.</u> Operator No. 3 and No. 4 stations are replicas of the two tactics stations in the aircraft. All aircraft tactical systems are physically and functionally replicated.

<u>Instructor's Compartment.</u> The instructor's compartment occupies the area between the flight trainees' compartment (aft end of the trailer) and tactic trainees' compartment (fore end of the trailer) and contains the flight and tactic instructors' stations.

The flight instructor's station contains the main console and the radio aids cabinet. This station contains all the facilities needed by the flight instructor to monitor all phases of flight. In addition to the normal complement of cockpit repeater instruments, controls are provided to modify flight conditions and induce a wide variety of aircraft system malfunctions. Adjacent to the main console is the radio aids cabinet; these systems were not used in either trainer observed in the course of the study.

The tactics instructor's station, with which we were most concerned in this study, consists of three consoles containing all the facilities required to introduce training problems, control target activity, introduce malfunctions into the various sensor systems, and monitor trainee performance. Individual panels are provided to load selected sonobuoys, insert target signal characteristics, and monitor operations of the ECM, JULIE, MAD, Radar, and ASN-30 equipments. Repeaters at this station are slaved to and provide a permanent record of MAD and JEZEBEL traces. Two closed circuit TV monitors permit operation of the ASN-30 equipment used by the TACCO to set up tactics to be monitored. A READOUT SELECT panel provides a means for obtaining a visual readout of the distance between any two selected objects (e.g., two sonobuoys, the aircraft and the target) involved in the tactics problem. A surface track recorder (plotter) is provided to present a graphic representation of progress through a problem. All of these displays were used by data collectors to compare actual data inputs (e.g., bearing and ranges) with those verbally reported by subjects.

VARIABLES AFFECTING CREW PERFORMANCE. This section discusses those variables in the training situation which could conceivably impact upon the performance of the S-2E crew. These variables are a function of the equipment designed for the mission (e.g., quality of the radar display) and of the nature of the mission itself (e.g., number of submarines in the search area). Some of the variables can be deliberately varied as an aid to learning (e.g., the number of submarines) while others, like radar display quality, cannot be varied (except in a special case in which the instructor causes the entire subsystem to malfunction).

The reason for considering this topic is that the variables, whether or not deliberately manipulated by the instructor, complicate the simulated mission and hence increase the difficulty of performing that mission. When the instructor deliberately utilizes a training input variable (by failing a sonobuoy, for example), for the express purpose of influencing the students' learning, that variable is in effect being used as an adaptive training variable. It is variables of this type that we are particularly interested in. Since student performance is possibly influenced by difficulty level, the reader must understand how the variables might affect the performance of our subjects. This can best be shown on a position-by-position basis.

Pilot. Although the number of variables producing difficulty (e.g., failures or malfunctions of navigation equipment, loss of an engine) that can be introduced is extremely large, the fact that we are not evaluating student performance in this position makes it unnecessary to describe them.

TACCO. Most difficulty variables having an effect on TACCO performance are second-order in nature; that is, they are variables which first affect the ability of the pilot, No. 3, or No. 4 operator to perform correctly. These effects in turn influence the TACCO's accuracy and speed of performance. For example, if the No. 4 operator reports an incorrect bearing or range, the tactical plot made by the TACCO with the ASN-30 equipment will be incorrect. In that respect the TACCO is primarily a recipient and processor of information from the other three crew members and is influenced by the accuracy of their inputs.

The only difficulty variable which acts directly on the TACCO is the failure of his ASN-30 equipment, either partially or completely. In the course of the study no such failure was introduced during data collection on TACCO trainees. Had these failures occurred, the TACCO would have been required to continue, using manual plotting and computation. This would have significantly affected response speed, but should have had no effect on accuracy.

No. 3 Operator. The non-acoustic operator is primarily responsible for monitoring three systems, radar, ECM, and MAD, each of which is affected by different variables:

ECM. Assuming proper operation of the ECM gear, the variables having the most significant effect on performance are the signal characteristics of the target. Variations in these characteristics (e.g., frequency, intensity, pulse width) produced by different targets can influence the classification performance of the operator. In addition to the signal characteristics, the number of targets and the complete or partial failure of the ECM are the only other failures of consequence. Since our subject sample for the No. 3 operator was small, the opportunity to observe the effect of these variables on ECM performance did not occur.

Radar. Assuming proper operation of the radar system, two variables would affect performance: range of the day, and number of targets. These variables would have independent and interactive effects on the operator's detection and classification performance. If the range of the day is restricted, targets outside that range would not be detected. Knowing the number of targets would make it easier to locate them geographically. In addition to these controllable variables, there are of course the whole host of display variables (e.g., resolution, distortion) which influence airborne radar system performance and which affect the operator's perceptual responses to stimuli. They affect his ability to monitor, detect, and classify targets presented on a PPI display.

MAD. Proper operation of the MAD system requires the operator to compensate for the magnetic fields present within the aircraft. Varying these factors either initially or at some point in the mission when MAD is active requires that the operator detect the change and take appropriate action. These factors were recorded as having been taken into account in the MAD operations observed. The only factor outside the trainees' control, which would tend to affect his performance was the effective range of the MAD system, which was adjustable by the instructor and would have implications for the response time available during MAD operations.

No. 4 Operator. The acoustic operator probably has the most important and demanding job in the system. In the majority of missions flown the performance of the TACCO and the pilot are directly dependent on the No. 4 operator's performance. As with the No. 3 operator, the variables impacting on his performance stem from the characteristics of the systems he utilizes. The two systems he employs and the factors affecting their performance are described below:

JEZEBEL. Assuming proper operation of the sonobuoy receiver system and JEZEBEL recorder system, the signals received from sonobuoy(s) are displayed to and analyzed by the No. 4 operator to determine the presence or absence of targets in the area and the nature of those targets. Performance of the JEZEBEL system is obviously contingent on the proper functioning of the sonobuoys in use and the sonobuoy receiver system; malfunctions in either of these two systems impose what is essentially a "no-go" condition on JEZEBEL recorder utilization. In addition to these go/no-go states, system functioning and operator performance are affected by buoy drift, sea state, ambient noise, range of the day, and partial malfunctions of an individual sonobuoy, all of which could be varied by the instructor.

JULIE. As with the JEZEBEL system, proper functioning of JULIE is dependent upon proper recorder; receiver, and sonobuoy functioning. The operator's ability to detect, monitor, and analyze signals displayed on the JULIE recorder is influenced by the same variables listed for JEZEBEL; in addition, the accuracy of his recorder functioning, the type of maneuver selected by the TACCO, and the accuracy of sonobuoy placement all affect the operator's ability to interpret JULIE signals.

Of the variables discussed only a certain number are controlled and varied by the instructor, although the trainer permits many more. In the remainder of the report reference is made to the selected list of variables actually utilized by the instructor with the term "training inputs." Note that the variables may change in amplitude (magnitude of the change) and frequency (number of times the variable occurs during a mission phase).

a. Multiple Targets. The presence of more than one target, while beneficial in allowing greater accuracy in constructing geometric fixes, generally tends to degrade the target detection and classification capability of each operator. A larger number of targets complicates the magnitude and complexity of the TACCO's tactical problem because he must receive and process more data and must make more complex decisions (e.g., which target to attack).

b. Target Speed Alteration. A decrease in target speed to slow or stop can reduce the No. 4 operator's ability to either detect or monitor the target's presence during JEZEBEL operations. Increases in target speed, on the other hand, enhance the probability of detection. Depending on the direction of speed changes and the effect on No. 4 operator performance, the TACCO's performance in computing target location is also affected.

c. Target Course Alteration. Changes in target course have an effect on the ability of both operators to detect and maintain contact with the target. Unnoticed or frequent changes in course make it extremely difficult for the TACCO to plot the course of the target.

d. Target Depth Changes. The effect of changes in depth depends on the particular depth selected and the stage in the mission. Operation of the submarine below periscope/transmitting depth renders the non-acoustic detection systems aboard the aircraft ineffective. The acoustical properties of the ocean environment are inconstant as to the enhancing or degrading effect on acoustic detector performance and accuracy (e.g., the relation of the target to a thermocline or occlusion). The result is that fragmentary data are supplied to the TACCO by the No. 4 operator; this in turn affects the TACCO's ability to compute target course and location.

e. Loss of Contact. The loss of target contact induced by the instructor in either the acoustic or non-acoustic mode requires that the crew (particularly the TACCO) approximate target location and revert to a search phase. Not only is mission time lost but also the target could remain undetected.

f. Recorder Failure. Failure of the visual display and recording capabilities of the acoustic system degrades the system and places additional burdens on the No. 4 operator. He now has only auditory stimuli with which to perform the same tasks and computations he could perform previously with visual and auditory displays. Failure of the MAD recorder does not impose the same burden on the No. 3 operator.

g. Sonobuoy Failure. The main effect of a sonobuoy failure is that no stimuli are received when the buoy is dropped. In addition to the need to replace the "dud" with another buoy, the failure may disturb the search pattern and require development of a new pattern. This results in lost time and difficulty in computing the nagivational and tactical position of the aircraft and target. Conceivably one could also prematurely expend available sonobuoys and thus preclude successful completion of a mission.

h. Receiver Failure. Failure of the receiver portion of any of the systems in the aircraft leaves the operator without a particular sensing capability (e. g., if a sonobuoy receiver fails, the No. 4 operator's listening capability is reduced; if the radar receiver fails, then the No 3 operator is blind to that spectrum, etc.). The implications of the various receiver failures are peculiar to each type of system; the amount of redundancy and the availability of backup capability vary as a function of the system, mission phase, and type of contact being prosecuted.

TRAINING DEVICE UTILIZATION. This section describes the way in which Device 2F66A was observed to have been used during the data collection period. Since the utilization pattern differed somewhat at the two training facilities, specific comments will be made to identify policy peculiar to either San Diego or Key West.

At both locations, the training device was scheduled and manned for operation on a 2-shift per day basis. The double shift operation was designed to maximize the availability of the device for the two major user populations: students and operational personnel (reservists constituted a minor subgroup of the operational population). In both San Diego and Key West slightly more than half of the available training device time was allocated to operational crews, with the remainder assigned to the training squadrons and their students. Two hours were scheduled for each training session.

Use of the simulator by operational squadrons was characterized by inconsistency in scheduling, crew assignment, mission selection, and overall performance. Adherence to the weekly schedule varied with the squadron, crew assignment, and training facility; as a result, a sizeable number of hours were unused by this population. The frequency with

which operational personnel appeared for training on the device seemed randomly controlled by overriding requirements of the crews' parent squadron. The missions, functions, and tasks practiced by operational crews when in the trainer were under the complete control of the crew commander, but did tend to follow the sequence of common mission phases identified earlier. In certain instances specific phases (JULIE, MAD, etc.) were singled out and practiced repeatedly.

The squadrons responsible for student training (VS-41 in San Diego, VS-30 in Key West) tended to utilize the afternoons or evenings of the trainer's operational day for training their students. Each location had a syllabus for the student training program, but as the details of each syllabus lesson were rarely adhered to it will not be described further here. Although the device was intended as a team trainer, especially in the later stages of training, it was generally the exception rather than the rule for the student population. In most of the training exercises observed, air crew training was carried out with the instructor assuming the roles of missing position personnel as necessary. In the more advanced training sessions at Key West, cadre personnel were available to operate vacant positions for training purposes. When personnel for all positions happened to be at the same training stage during the same time frame and their progress in the WST syllabus roughly coincided, student crews were scheduled for training together.

One difference in S-2E training at Key West lay in the fact that an additional training device, an S-2E Aids Trainer (Device 14B30), was used intermittently for training of the No. 4 operator. Instructor personnel at Key West expressed a preference for using this device because it allowed them to work alongside the student rather than being physically separated as they were in the 2F66A.

Inflight training was carried on concurrently with WST training at Key West; although similarly scheduled in San Diego, it was actually begun there extremely late in the syllabus (and in some cases after all WST was completed).

SUBJECT POOL

Considering the relatively small number of personnel undergoing S-2E training, it was unnecessary to select subjects; all personnel in training at VS-41 in San Diego were designated part of the subject population. As the study progressed, it became apparent that even this pool of student subjects did not constitute a large enough sample; consequently an additional data collection effort was undertaken at VS-30 at Key West.

A definition of the "subject population" is presented below:

STUDENT PERSONNEL. Student personnel consisted of officers (TACCOs) and enlisted personnel (No. 3 and No. 4 operators) being trained by the Replacement Air Group (RAG) for duty in operational VS squadrons. The

personnel observed during data collection were receiving initial training in the aircraft type after having completed required academic training at either FAETUPAC or FAETULANT.

OPERATIONAL PERSONNEL. The operatio. ' personnel were crews from squadrons other than VS-41. Such personnel already had at least one operational cruise in S-2E aircraft. Operational crews did not train under the auspices of the RAG, but rather were under the control of their individual squadrons. As indicated previously, these crews did not report to the device for training on a systematic basis. The implication of this for the data collection effort was that whereas certain crews were observed in the trainer on several occasions, others appeared only once and of course not every operational crew was observed.

RESERVE PERSONNEL. Reserve personnel were members of Naval Reserve Squadron VS-68R on-board at San Diego for two weeks' annual training, during data collection.

The personnel constituting the subject pool must be separated not only by population, but also by position: TACCO No. 3 operator, and No. 4 operator. The training situation and the position specificity of the tasks necessitated the development of performance measures identified by, and uniquely for, each position rather than for a team.

Table 1 presents the total number of personnel identified by population and crew position serving as subjects for this study at the two training facilities. Several of the subjects provided repeated measures, i.e., were observed for more than one training session. Tables describing the subject populations per session were especially pertinent to the data analysis discussion and, hence, are presented in that section.

The reader should be warned that the numbers in Table 1 do not correspond to the numbers of subjects actually available at any individual training session (e.g., session 4) or to the number of personnel represented by data points in the learning curves presented later. The reason is that a larger number of personnel were available than ever served as subjects at any one time. Specifically, personnel dropped in and out of the pool for the following reasons:

a. A few students were already in training at the time data collection began, and completed their training after being subjects for a few advanced sessions.

b. Some students, particularly operational personnel, left the training program for one reason or another before data collection was completed. These subjects too are represented only in a few training sessions.

TABLE 1. THE SUBJECT POOL

SAN DIEGO

	TACCO	No. 3 Operator	No. 4 Operator
Students	3	1	9
Operational*	31	33	29
Reserve**	16	15	15

*27 teams plus 7 No. 3 operators with no team. Higher numbers result from crew members performing in other positions on different sessions. (NOTE: All 27 teams were not complete. The TACCO and pilot were never observed for 2 of the 27 teams; neither was the No. 3 operator on one other team.)

**Number of teams equals 14. Higher numbers result from crew members performing in different positions on different sessions.

KEY WEST

	TACCO	No. 4 Operator
Students	6	6

c. A few students entered the training program after data collection was initiated and were subjects for beginning sessions only.

Ordinarily, if a large enough pool of subjects had been available personnel falling into the previous categories would not have been accepted for the subject pool. However, considering the restricted N available for the study, it was considered wiser to include them for the training sessions in which they participated. Descriptions of the subject pools by session are presented in Tables 4 and 5 (see pp 32-33).

DATA COLLECTION METHODS

DATA COLLECTION PERSONNEL. The data reported in this study were collected by two Bunker-Ramo human factors personnel who were present in the training device throughout each training day of the evaluation period. The device was scheduled for training from 0800 to 2200 each working day. The data collectors worked on a two shift basis—one from 0800 to 1500, the other from 1400 to 2200, allowing data collection to proceed during all training sessions.

(One might ask whether two investigators during a training session might not supply more data than one would. Because of the size of the trailer in which the device was located, only one data collector could be present at a time. Moreover, it was ascertained empirically before formal data collection began that there was no substantial data loss with only a single data collector.)

The data collection personnel were not, of course, skilled operators for each position being trained. However, they had received a one week training course in the operation of the ASN-30 (Device 14B30) ordinarily given to new TACCO students prior to the start of their training on the 2F66A. This training gave them "hands-on" experience aboard the S-2E. They had also studied operational and training manuals for the device, observed training sessions prior to data collection, and assisted in the selection and development of the performance measures which were employed.

To recognize the constraints under which these personnel collected data, it is necessary to understand that all data were collected under non-interference conditions. In other words, a necessary condition for performance of the study was that data collectors could not modify or interfere with the manner in which training was conducted. However, they were allowed to ask questions of the instructional staff controlling the training inputs, and to talk to student personnel when the latter were free.

TYPES OF DATA COLLECTED. Before describing the measures in detail it will be helpful to the reader if we summarize the types of data collected during the training session. These were of four types:

(1) objective performance measures, (2) instructor evaluations of student performance, (3) instructor evaluations of the trainer, and (4) recordings of student communications during training sessions.

The complete set of objective performance measures are presented in Table 2, while Table 3 lists the objective performance measures found to be sensitive to training on the 2F66A. These are categorized by crew member position and mission stage. A detailed description of each measure and its rationale follows Table 3.

Two types of instructor ratings of student performance were available. These were: (1) evaluation of the student's performance in a particular training session, and (2) evaluation of the improvement in performance over the preceding training session, for both the overall crew and the individual crew member.

The evaluation in the initial training session was done as follows:

a. In terms of overall crew performance, using a 6 point scale ranging from minimal to above average proficiency

b. In terms of the individual crew position, utilizing the above average (AA), average (A), and below average (BA) scale which instructors customarily use in evaluating student performance.

Appendix B presents the rating package which instructors were asked to fill out at the conclusion of each training session.

Instructor evaluation of the trainer were made subsequent to data collection, at the conclusion of the study period. These evaluations were secured at both San Diego and Key West. Sound recordings of crew communications were made for each training session. Since the training session ran approximately two hours, it was necessary to sample the communications at the rate of five minutes out of every twenty.

<u>Objective Performance Measures</u>. The performance measures used during this study were measures derived from and based upon the operational missions of the S-2E aircraft. These measures were derived from an intensive analysis of the S-2E ASW mission and the roles of the individual crew members in the performance of that mission. The results of this analysis, and the measures that it implied, were submitted for review to the instructional staff of VS-41. Based upon the suggestions of the instructional staff the measurement set used during the study was revised prior to initiation of data collection.

The performance measures used can be categorized as either <u>terminal</u> or <u>intermediate</u>. Terminal measures are those which describe functions and tasks representing mission completion. For example, in the Air

TABLE 2. PERFORMANCE MEASURES FOR WHICH DATA WERE COLLECTED

Position	Measure
TACCO	1. Range and bearing position errors for sequential buoys during LOFAR (actual versus logged position) 2. Range and bearing errors for target during CODAR 3. Bearing error for buoy drop points during CODAR 4. Appropriateness of JULIE pattern selected 5. Accuracy of pattern (range and bearing deviations from doctrine) 6. Number of kills attempted 7. Number of kills 8. Miss distance
No. 3 Operator	*1A. Radar detection performance (time to detect) 1B. ECM detection performance (time to detect) 2A. Radar classification performance (accuracy of classification) 2B. ECM classification performance (accuracy of classification) 3A. Radar range and bearing error (position accuracy) 3B. ECM range error (position accuracy) 4. Time between updates (radar or ECM) 5. On Top Time (computation based on distance, A/C speed) 6. MAD detection performance (probability of detection) *The performance of some functionally identical tasks by the No. 3 operator were different due to the equipment used. 1A identifies a task performed with radar. 1B identifies the same task performed with ECM gear.

TABLE 2. PERFORMANCE MEASURES FOR WHICH DATA WERE COLLECTED (continued)

Position	Measure
No. 4 Operator	1. Time to detection during LOFAR 2. Classification time during LOFAR 3. Classification accuracy during LOFAR 4. Time to detection during CODAR 5. Time between updates 6. Range error to target during JULIE 7. Number of missed buoy echoes during JULIE 8. Calibration range error during JULIE

TABLE 3. LIST OF OBJECTIVE PERFORMANCE MEASURES FOUND TO BE LEARNING SENSITIVE (BY POSITION)

Position	Letter	Mission Phase	Measure
TACCO	A	CODAR	Buoy drop points, position with relation to bearing, error in degrees
	B	JULIE	Pattern accuracy with relation to range, error in yards
No. 3 Operator	A	LOFAR and/or CODAR	Range to target, error in miles
	B		Time between updates, minutes
	C	MAD	Target detection probability (number detected divided by number of opportunities for detection)
No. 4 Operator	A	LOFAR	Time to classify target, minutes
	B	LOFAR	Proportion of accurate classifications
	C	JULIE	Range to target error, in yards

Antisubmarine Warfare (AASW) mission the end result of the tactical operations is the release of weapons against a target. Two types of terminal measures result: whether or not the target was "killed," and the distance from the target at which the weapon detonated ("miss distance").

Intermediate measures describe the operations which lead up to or implement the completion of the mission, but do not themselves describe mission completion. For example, the placement of sonobuoys in a LOFAR or CODAR pattern, listening to returns, and detection of a contact are all intermediate mission activities from which intermediate measures can be secured.

Intermediate measures are primarily measures of individual position performance; terminal measures reflect crew performance. Almost all of the objective performance measures secured in this study were intermediate measures. This presents some problems in determining the effectiveness of crew functioning as a whole. If the crew does not proceed to the terminal activity of the mission (as was most often the case in the present study), it is difficult to specify crew efficiency for that mission and session.

Table 2 contains a listing of the performance measures set for which data was collected during the course of this study; during every observed training session efforts were made to secure data for each of these measures. Table 3 represents those measures retained after data reduction and analysis; as can be seen by the difference between Tables 2 and 3, a large number of the measures initially proposed have been discarded. Those measures which have been retained are the result of screening each measure and its associated data against several criteria. These criteria included:

a. Sample Size. Measures for which data was either nonexistent (zero observations) or for which the sample size numbered two or less were rejected. An example of this is provided by the ECM measures for the No. 3 operator. ECM gear was so rarely utilized that practically no data resulted and no systematic training can be said to have occurred.

b. Nonpersonnel Sources of Variance. Those measures which, on the basis of data collection records and in depth interrogation of the data collectors, were found to be subject to factors other than personnel performance were rejected as measures of device performance. The source of the confounding variance differed from measure to measure but included such factors as instructors fixing a particular tactic, computer or device values which were frequently in error to some variable degree, and unsynchronized repeaters.

c. Valid Sensitivity to Learning. Those measures retained after screening against criteria (a) and (b) on the preceding page were then analyzed to determine if performance changes occurred which could be ascribed to continued trainer use rather than other variables (such as intervening flights and training inputs). These analyses are described in greater detail in the data analysis section.

Only those measures which successfully met the criteria are presented in Table 3 and in the results section of this report. It should be noted that in a complex training task, the investigator must either measure literally "everything that moves" or have some rationale for having selected the measures he has chosen. In the present study the primary rationale was the need to define measures relevant to operational criteria directly related to operational objectives. Support for use of some of the measures selected was found in a similar study performed by Metersky (1967) which investigated the effects of fatigue on individual and crew performance in an S-2E simulator. Although Metersky used a smaller measurement set, fatigue effects and skill level effects were found for a few of the measures. In summary, the primary rationale for the initial measurement set was its relevance to operational performance requirements, supported by evidence that performance level differences did exist as a result of skill level.

For all the methodological complexity of these measures, the means required to record them were quite simple. Stop watches calibrated in tenths of seconds were sufficiently precise to record time differentials. This is because fractions of seconds are apparently not critical to performance of the mission, at least not within the trainer. Errors could be recorded by visual and aural observation, since the instructor's station possessed a number of repeaters replicating student displays and readouts, which presented the computer's calculation of actual bearings and ranges. When the determination of correct performance depended on instructor judgment, this could be recorded by monitoring the instructor's comments or by questioning him during the training session.

Listed in subsequent paragraphs is a detailed description of those objective performance measures found to be sensitive and reliable indicators of performance for the mission phases, functions, and individual tasks trained on Device 2F66A. It should be noted that in the majority of cases the particular behavior described by an individual measure occurred more than once in a session (e.g., each of the several buoys dropped during CODAR required a position report, each of the targets presented during LOFAR required classification) thereby allowing a reliable measure to be obtained. The individual session time and error measures for a subject are mean scores, while accuracy and probability measures are expressed as proportions of the total.

TACCO

a. Measure A: Bearing error in position of buoy drop points during CODAR. As the TACCO makes successive buoy drops, he reports and logs the bearing of each of those drops. This reported bearing position was compared by the data collector with the computer readout of buoy bearing for each successive buoy drop point.

b. Measure B: Range error in buoy pattern accuracy during JULIE. During JULIE as the aircraft flies a preselected pattern, buoys are dropped at predetermined ranges and bearings (these last according to the requirements of the selected pattern). As the TACCO drops each buoy, he announces the range of each drop from the previous buoy. The actual range of the drop from the previous buoy is read out at the instructor's station. The difference between the reported drop point and the actual position is the range error. The mean of successive range errors for each subject is what is being dealt with.

No. 3 Operator

a. Measure A: Range error to target during LOFAR/CODAR. This measure is a comparison of the actual range of the target from the aircraft when detected on radar versus the range called out by the No. 3 operator. The concern here is with the operator's perceptual error in interpreting the range data on his scope. Variations in the actual range of the target from the aircraft would not significantly influence perceived range since the target would always be visible on the scope when range was reported. Actual target range was determined from readouts at the instructor's station.

b. Measure B: Time between information updates during LOFAR/CODAR. This measure is the duration between the operator's previous verbal report and the one being recorded. A number of factors influence this measure: (1) operational procedure, which dictates the frequency of update reports, and (2) distance of the aircraft from the target, which is also affected by aircraft speed. These two factors obviously interact. Procedure requires that when range to target is over ten miles, a report must be made every two miles. Under ten miles but over two, a report is made each mile. Under two miles, a report is made every quarter mile; since the procedure is fairly specific, deviations can be considered to be measures of a person's ability to "keep on top" of the situation, monitor time and speed, and time share these with any other tasks he may be performing.

c. Measure C: Probability of target detection during MAD. This is a ratio between the number of targets detected and the number of targets available during MAD for detection. The higher the ratio (the more it approaches 1.0), the better the performance is.

No. 4 Operator

a. Measure A: Time to classify target in minutes, during LOFAR. This measure represents the time between the operator's detection of the target and the time at which he reports a classification of the target as being submarine or nonsubmarine and its type. This measure follows immediately—or relatively so—upon target detection. Since the data collector has recorded the time at which the target was detected, he merely follows this by recording the time when classification was made. Since sea state was a constant during the problem, this variable did not affect classification speed. Nor would range and bearing of the target affect classification speed significantly.

b. Measure B: Classification correctness, during LOFAR. Classification accuracy was determined by comparing the verbal description of the target with the known characteristics of the target inserted into the problem. This measure was a binary one, with "0" given for correct classification, "1" for incorrect classification. Because of problems with realistically simulating target characteristics, instructors did not stress classification, and partial values for various target characteristics could not be assigned to the operator's responses.

c. Measure C: Range error to target, in yards, during JULIE. During JULIE operations, the operator is required to make a verbal report of the range from the target to a particular buoy, based on his interpretation of his range recorder traces. To secure this measure, a comparison of the verbal report in yards was made with a computer readout of the actual target range. The difference represents the range error.

Instructor Evaluations of Student Performance. At the conclusion of each training session the instructor was asked to rate overall crew performance on a 6-point scale, ranging from minimal to above/average proficiency. He was told that he could place his X at any point along the following scale continuum.

Minimal Proficiency					Above Average Proficiency
'	'	'	'	'	'

Following this he rated the three positions individually in the order of TACCO, No. 3, and No. 4 operators. The rating was to be based on the instructor's perception of how well the student had performed in relation to the objective performance measures described previously. The crew member was first rated on the AA, A, BA scale customarily employed. Then the instructor was asked to indicate, within the AA, A, BA scale category selected, how well the student had performed. This was a 9-point scale ranging from low to high. In other words, assuming that a crew member had been rated as average (A) in performance, the instructor could then indicate whether he was a high average, intermediate average, or low average.

In the event that the instructor evaluating the crew had observed their performance on a previous session (which unfortunately did not happen as often as desired because of the variability in instructor scheduling) he was asked to indicate (again on a 9-point scale—low to high) their proficiency on the preceding training session in relation to the present session. This indicated the extent of their improvement, if any. The same procedure was followed in instances where only an individual crew member was undergoing training; however, on these occasions ratings of crew performance were not required.

Instructor Evaluations of Trainer Characteristics. At the conclusion of the study, instructors at both San Diego and Key West were asked to evaluate training device characteristics in terms of the individual crew station (pilot, TACCO, No. 3 and No. 4 operators) with which they were most familiar. (Appendix B presents the instructor rating package for trainer characteristics.)

Communications Recordings. The objective performance measures obtained reflect primarily individual position performance, even when recorded in a team training context. Since the trainer was developed primarily to be a crew training device, it was desirable to secure some measures which reflected crew interaction exclusively. Since that interaction was performed through communications, it was felt that the analysis of crew communications would indicate how crew interaction was affected by training. In other words, if certain communication indices were assumed to represent crew interaction behavior, and these indices showed a change in relative frequency of occurrence as a function of training, it could be presumed that training sessions had influenced crew interaction.

The specific communications measures are described in the paragraphs to follow. In general, questions (about course of action, information, etc.) were asked by one crew member and were responded to by another. Items of information of one type or another were volunteered. Commands were issued and acknowledged; irrelevant statements were made. Since verbal exchange between the two operators and the TACCO is a basic requirement

for mission performance, it can be assumed that, if the frequency of relevant communications and the direction of communications change meaningfully over the training sessions, crew coordination has increased.

Measurement of communications activities was obtained with samples of communications during several of the two hour training sessions. Originally it had been intended to sample five minutes at the start and end of each mission phase. The sampling procedure was modified, however, when it became apparent that (1) mission phases were often repeated during the same training session; or that (2) mission phases being simulated were not always completed. The modified recording rate was five minutes out of every twenty throughout the session.

A sampling procedure was utilized instead of a complete session recording, because of the length of the session and the consequent amount of material which would result. Comparison of data taken from different sessions, under any one condition, indicated that the sampling procedure produced data that was reliable. The number of student, operational crew, and reservist sessions for which communications were analyzed were: students, 37; operational personnel, 10; reservists, 11.

The procedure used in analyzing these recordings to develop a data set was comparatively straightforward. The recorded material was listened to and sorted into a set of categories developed on the basis of analysis of categories used in previous studies, particularly those of Krumm and Farina (1962) and Briggs and Johnston (1967). The categories used by Federman and Siegel (1965) were also reviewed, but were not used because they required excessively subjective judgments on the part of analysts.

The categories finally selected are listed below, along with examples that will help define them:

a. Requests Information, Opinion or Course of Action. Example: "What was the time on that last drop?" "What do you want to drop next. . . North?" In the original Krumm and Farina categories, information and opinion were subdivided, but it was found during preliminary analysis that too fine a line had to be drawn between opinion and factual information. Originally, course of action was subdivided from information or opinion, but the frequency of the former was so low that for statistical treatment it had to be combined with the latter. (Identified as C1 in Results.)

b. Provides Requested Information, Opinion or Course of Action. Example: "The time on that drop was 3:8." "No, come around to port to drop Maypole 23 to south." Response to (a) above. (Identified as C2 in Results.)

c. Requests Permission. Example: "Request permission to lower radome and MAD boom." This category was developed specifically for this study; however, no instances of this type of communication were recorded.

d. Provides Disposition of Request. Example: "Permission granted." Response to (c) above; again, however, there were no occurrences of this category.

e. Declarative Information (Including Direct Commands). Example: "Maypole 17 away now, now, now," or "Suggest heading of 090." The essential characteristic of this category is that it is a voluntary input, rather than a request or response to a request. The original categories from Krumm and Farina were "volunteers course of action" and "volunteers opinion." However, it was impossible to achieve a significant level of interrater agreement on the difference between course of action and opinion. Consequently the category used by Briggs and Johnston was employed. Because of the low frequency of the direct command category, this category was eventually combined with declarative information. (Identified as C3 in Results.)

f. Verifies Information (Feedback). Example: "Roger, understand 090," or "17 indicates away." This category is also a new one, not found in either Krumm and Farina or Briggs and Johnston. Nevertheless, it was frequently found in the recordings of the present study. (Identified as C4 in Results.)

g. Acknowledgements. The Krumm and Farina category of "formal indication of compliance to orders" was too easily confused with the "Acknowledgement" category; hence, the former was eliminated. Typically, a response of this nature consists merely of the statement "Roger," "OK," which is more easily handled under the acknowledgement category. Response to (e) above. (Identified as C5 in Results.)

h. Irrelevant Remarks. Is the same as the Krumm and Farina category. These are comments which have no relationship to the completion of the mission phase or immediate problem, such as "This thing's all screwed up." "B_S_." (Identified as C6 in Results.)

DATA ANALYSIS METHODS

The nature of the data analyses conducted for this study, the manner in which the analytic results were reported, and the interpretations made from the results have been directed by the characteristics of the collected data.

Field studies suffer from a type of problem diametrically opposed to that of laboratory investigations of human behavior. Whereas the laboratory method permits the controlled manipulation of variables and precision of measurement, control is often obtained at the expense of a realistic representation of actual conditions. Conversely, in field evaluations, the desired amount of control is usually lacking, but the realism of the training situation is there with all its confounding variables.

The "not-to-interfere with training" basis under which the evaluation was conducted is a handicap to controlled experimentation. Too much control, however, can affect the realism of the training situation and thereby create an artificiality that can produce misleading results in the evaluation. One purpose of the early studies is to analyze the amount of experimenter control which is necessary but which will not contaminate the results, and in so doing, to develop practices that will increase the validity and reliability of field evaluations.

The requirement to collect performance data in a training situation both purposely flexible (to meet individual student needs as perceived by the instructor) and subject to random interference (e. g., equipment and computer malfunctions, personnel scheduling), and for which the number of students in any one class is very small, results in data which can be very difficult to effectively analyze because:

a. The lack of experimental controls allows a multiplicity of variables to operate concurrently.

b. The small sample size allows only a limited number of variables to be evaluated in any one analysis.

c. Lack of controls combined with a small sample size can result in variables confounded so that their separate effects on performance are impossible to determine.

Tables 4 and 5, which describe the training sessions observed by population, crew position, and session, provide an illustrative indication of these problems. As the tasks are mission-phase restricted, the amount of objective performance measurement data was further reduced by the fact that not all subjects listed for a particular session in Table 4 or 5 necessarily performed in all mission phases during that session. The characteristics of the data collected are such that the utilization of inferential statistics would have had to be based on numerous assumptions and subjective judgments that could not have been substantiated. Interpretations rendered on the basis of measures of significance would be very questionable, if not spurious. Therefore, in the interest of providing results based on objective procedures, clearly stated assumptions, and analytic techniques appropriate to the data, descriptive statistics have been the primary analytic tool. (The only exception to this is the retention of the chi-square statistic for the communication analyses. In

TABLE 4. THE STUDENT POPULATION DESCRIBED BY THE INDIVIDUAL STUDENTS, THE LOCATION OF THEIR TRAINING, THEIR POSITION, AND THE SESSIONS THEY PERFORMED DURING DATA COLLECTION PERIOD

	Session Number								
Student Identifier	1	2	3	4	5	6	7	8	9
<u>TACCO Students</u>									
San Diego:									
A	X								
B				X	X	X	X		
C				X	X	X			
Florida:									
D				X	X				
E				X	X				
F					X				
G				X	X				
H				X	X				
I				X	X				
<u>No. 3 Operator Student</u>									
San Diego									
J			X	X	X	X			
<u>No. 4 Operator Students</u>									
San Diego:									
K	X	X							
L	X	X	X	X	X	X			
M	X	X	X	X	X	X	X	X	
N	X								
O	X	X	X	X	X	X	X	X	
P			X	X	X	X	X	X	X
Q			X	X	X	X	X	X	
R			X	X	X	X	X	X	X
S			X	X	X	X	X	X	
Florida:									
T						X	X	X	X
U						X	X	X	
V							X	X	X
W							X	X	
X							X	X	
Y							X	X	X

TABLE 5. THE RESERVE AND OPERATIONAL POPULATIONS DESCRIBED BY POSITION AND BY SESSION. . OBSERVATIONS WERE MADE ONLY IN SAN DIEGO AND INCLUDED ALL RESERVIST AND OPERATIONAL PERSONNEL TRAINING SESSIONS

Reserve Population

	Session Number		
Position	1	2	3
TACCO	16	9	3
No. 3 Operator	15	9	4
No. 4 Operator	15	8	4

Operational Population

	Session Number				
Position	1	2	3	4	5
TACCO	31	8	3	1	
No. 3 Operator	33	9	2		
No. 4 Operator	29	11	4	1	1

this instance, the measures and the hypotheses not only met the chi-square assumptions but further, provided unusually reliable data.) The deductive procedures for exploratory investigation were followed: general hypotheses are examined first, with subsequent examinations of more detailed hypotheses serving to validate (or invalidate) and more usefully define the generalized results.

Because the primary goal of the study was to determine the effectiveness of the trainer in improving crew member performance, the data has been separated by training session to demonstrate the effects of continued device use. The definition of session number (using session number 2 as an example) is described for each population as follows:

a. Student: Session number 2 is the second time the student has actually performed in the trainer.

b. Reserve: The second time the reservist has performed in the trainer during the annual tour of active duty.

c. Operational: The second time that a member of an operational squadron performed in the trainer during the data collection period.

Based on the above considerations, the data presented in the Results section of this report consist of: (1) learning curves plotted for those objective performance measures found to be clearly interpretable in terms of trainer effectiveness (i.e., where reasonably consistent performance changes as a result of practice are clearly evident), (2) learning curves plotted for the subjective measures, i.e., instructor ratings of trainee performance, (3) measures of communications performance, and (4) instructor evaluations of trainer effectiveness. The data analysis methods utilized for each of these are discussed in greater detail in the following paragraphs.

LEARNING CURVES FOR OBJECTIVE PERFORMANCE MEASURES. The most descriptive method of presenting the results of research into training and learning performance is by means of the learning curve. The learning curves presented in the Results section of this report are based on both objective (i.e., quantitative operationally relevant performance) and subjective (i.e., instructor evaluation of personnel performance) measures for each crew position. Those measures which were interpretable in terms of trainer effectiveness are presented in one or both of two forms: generalized learning curves and population learning curves. These curves are presented on the basis of the following assumptions and procedural rules.

Generalized Learning Curves. The generalized learning curves present measurement data that has been separated by observed training session. The generalized learning curves present data that has been pooled with

respect to: (1) populations (student, operational, reserve) and (2) location (Key West and San Diego). This pooling of data was based on the following assumptions: (1) the personnel observed at each position were a representative sample of Navy personnel who would perform these same tasks in the operational environment, and (2) the training devices upon which the subjects performed present an adequate representation of a simulated operational task in representative training environments. The generalized curves were developed for only the No. 4 operator position. The limited data available on the student TACCO and No. 3 operator populations did not permit the pooling of population data for these positions.

Population Learning Curves and Other Curves. Separate learning curves are presented for each of our populations to describe population performance differences as a function of continued trainer use. The effects of intervening flight and initial crew training sessions on the student population are also presented for two measures where the effects were relatively consistent and the data sample reasonably controlled with respect to other variables.

The generalized and population learning curves presented in the Results section are for those measures retained after assessment of all measures against these previously cited criteria: sample size, nonpersonnel sources of variance, and valid sensitivity to learning. The application of the first two criteria is obvious. The methodology related to the third criterion is, however, more complex and warrants further discussion. The data for each performance measure were separated according to session number and individual subject. The data were identified such that the effects of the following variables on performance could be defined: continued trainer use for each population, intervening flights, trainer inputs, trainer characteristics, and crew (as opposed to individual) training sessions. The question under consideration was this: Is the learning curve demonstrating effectiveness or ineffectiveness for Device 2F66A for training the measured behavior valid, or is the demonstration either artifactual or confounded by other concurrent events? The question was evaluated against each measure by graphic analysis of individual subject performance curves plotted across sessions. The curves were examined both individually and collectively to determine the existence of any consistent behavioral response to the above variables. Examples of the graphic analyses performed are provided by a discussion of Figures 13 and 18 in the results section. Figure 13 presents individual performance curves for the No. 4 operator student population on Measure 1: "Time to Detection During LOFAR." It should be noted that generalized and population learning curves are not presented for this measure, nor is the measure listed in Table 3 as "learning sensitive." The reason is evident upon examination of Figure 13: individual performance level variations were not relatable to the sequence of sessions; nor were they found to be relatable to the occurrence of any other variable given above except for one: the occurrence of the student's first crew training session.

Figure 18, on the other hand, presents a subset of the data for Measure C of No. 4 operator performance: Range to Target Error During JULIE. This measure was selected as a learning sensitive measure and the learning curves are presented in Figures 11-1 and 11-2. Examination of the performance data described by the individual student performance curves in Figure 18, however, indicate that student performance on this task may be affected by another variable as well: the intervention of flight training between trainer sessions. The graphic evidence was supported by the fact that with the exception of trainer sessions and intervening flights, all of the other indicated variables were found to be constant across this data subset.

LEARNING CURVES FOR SUBJECTIVE PERFORMANCE MEASURES. The ratings given to individual personnel by their instructors were averaged and plotted for each successive trainer session in the same manner as described above. The result is a learning curve like those previously described, but a curve based on subjective evaluations of performance, rather than on objective measures of performance.

INSTRUCTOR EVALUATIONS OF TRAINER EFFECTIVENESS. The evaluations are of two types: quantitative and qualitative. Quantitative evaluations were ratings of various trainer features. Analysis of these ratings consisted of developing means and ranges for each crew position and for each trainer feature being evaluated. The most commonly occurring qualitative responses were simply listed.

COMMUNICATIONS ANALYSIS. It will be recalled that the communications during most two hour training sessions were sampled and recorded. The basis of sampling was five minutes out of every twenty. It was hypothesized that crew coordination effects resulting from simulator practice would be manifested primarily through operational procedures which to an extent dictate a particular pattern of communication.

The first two tapes were analyzed by two independent evaluators to secure a measure of interrater reliability. It was assumed that if this reliability were high enough, the task of evaluating the subsequent tapes could be handled by one of the two raters. This at least was the procedure employed by Federman and Siegel (1965). Interrater reliability in the final categorization format was found to be .97, which was considered high enough to permit one man to continue the analysis. The analysis therefore, consisted of determining the frequency of (1) communications within each category and (2) interposition communication exchanges. The frequencies were transformed to percentages such that the changing pattern of category usage and of interposition exchange across sessions would be evident for each position. Chi-square analyses were performed on the frequency distributions to assess the significance of distribution differences.

SECTION II

STUDY RESULTS

INTRODUCTION

The results of the study will be described in terms of the five major study goals:

a. Do students trained on the device learn?

b. How does learning on the device compare among students, operational personnel, and reservists?

c. How is learning influenced by various training inputs?

d. How is learning influenced by the introduction of team training and by the sequencing of simulator and flight training sessions?

e. How do instructors evaluate trainer effectiveness?

DO STUDENTS TRAINED ON THE DEVICE LEARN?

The primary data bearing on this question consist of learning curves (generalized, population) which are broken out by individual crew member position (No. 4 operator, TACCO, No. 3 operator) and by individual system measure. The axes against which the learning curves are plotted are defined as follows: the x-axis, or abscissa, identifies point in training, i.e., session number; the y-axis identifies the performance level attained for a particular measure. The data items to be considered in answering the above question were summarized as follows:

a. TACCO

1. Measure A: Position of buoy drop points, CODAR

2. Measure B: Buoy pattern accuracy, JULIE

3 Measure C: Instructor ratings of TACCO performance

b. No. 3 Operator

1. Measure A: Range error to target

2. Measure B: Time between information updates

3. Measure C: Target detection probability, MAD

4. Measure D: Instructor ratings of No. 3 operator performance

c. No. 4 Operator

1. Measure A: Time to classify target, LOFAR

2. Measure B: Classification accuracy, LOFAR

3. Measure C: Range error to target, JULIE

4. Measure D: Instructor ratings of No. 4 operator performance

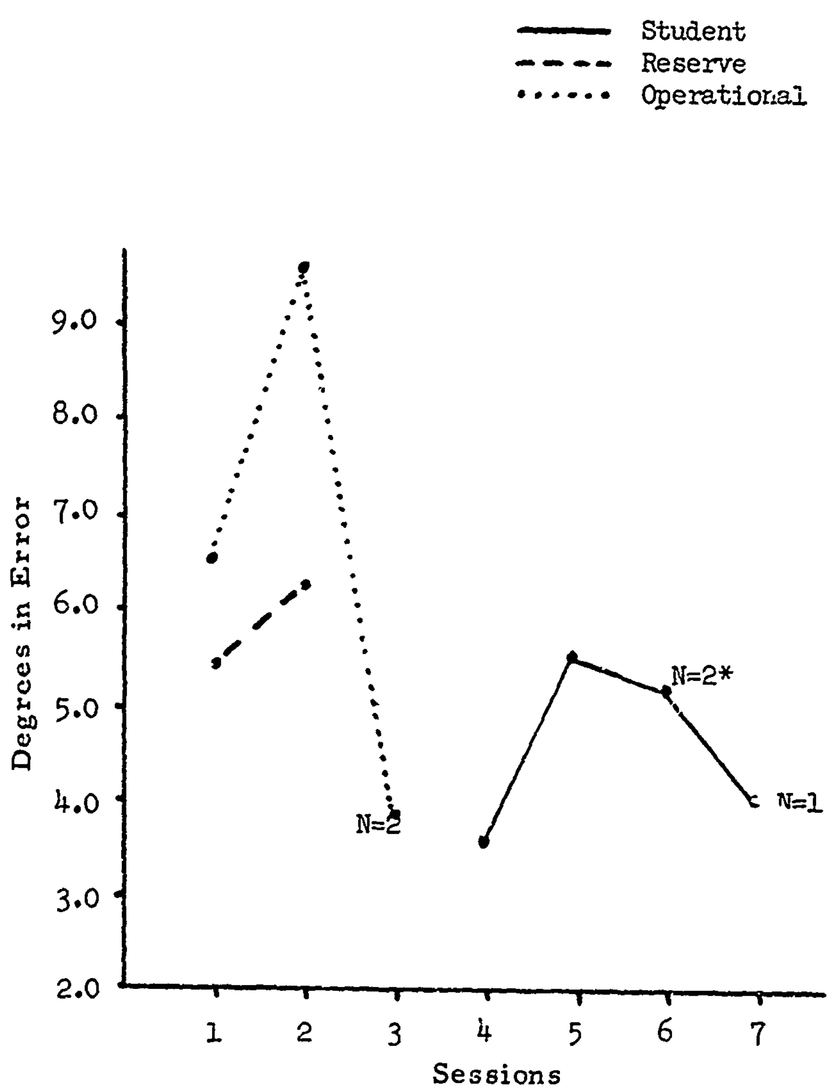
TACCO

Measure A: Bearing Error in Positional Buoy Drop Points During CODAR. This measure is a comparison between the TACCO's announced placement of successive buoy drops and the computer readout of actual drop point location.

The overall configuration for the population curves displays a negative tendency which indicates that performance on the task may be improved by training on this device. The Ns for individual data points were too small for validation analyses beyond the determination that continued use of the trainer tended to relate to improved performance. The population curves presented in Figure 2 are also erratic, a tendency which was characteristic of much of the TACCO data. The erraticism might be interpreted as implying that: (1) other factors (e.g., secondary task loading, etc.) interfere with or degrade performance on this task, and/or (2) the performance itself is not reliable to more than ±1 or 2 degrees, and/or (3) continued training beyond that observed would be required to improve or stabilize performance.

Measure B: Range Error in Buoy Pattern Accuracy During JULIE. In a given JULIE pattern buoys are supposed to be spaced at predetermined ranges and bearings. The TACCO, as he places each buoy, indicates the range of each drop point from the previous buoy. The actual buoy spacing is available from the computer. The difference between the actual and reported spacing is the range error.

Examination of the population curves presented in Figure 3 indicates that performance on this task by the student population was quite erratic. Contrasting student TACCO performance with that of operational and reserve TACCOs indicates that during the training sessions observed, student performance did not achieve either the accuracy or the consistency that can be expected of operational or reserve TACCOs. Limited evidence of benefit from trainer use is provided by the reserve population.



*Note: The population for a given data point is presented whenever that value is less than 4. An N of 4 was generally found to provide a stable estimate, while Ns of less than 4 did not.

Figure 2. TACCO, Measure A: Positional Error of Buoy Drop Points During CODAR, Population Curves

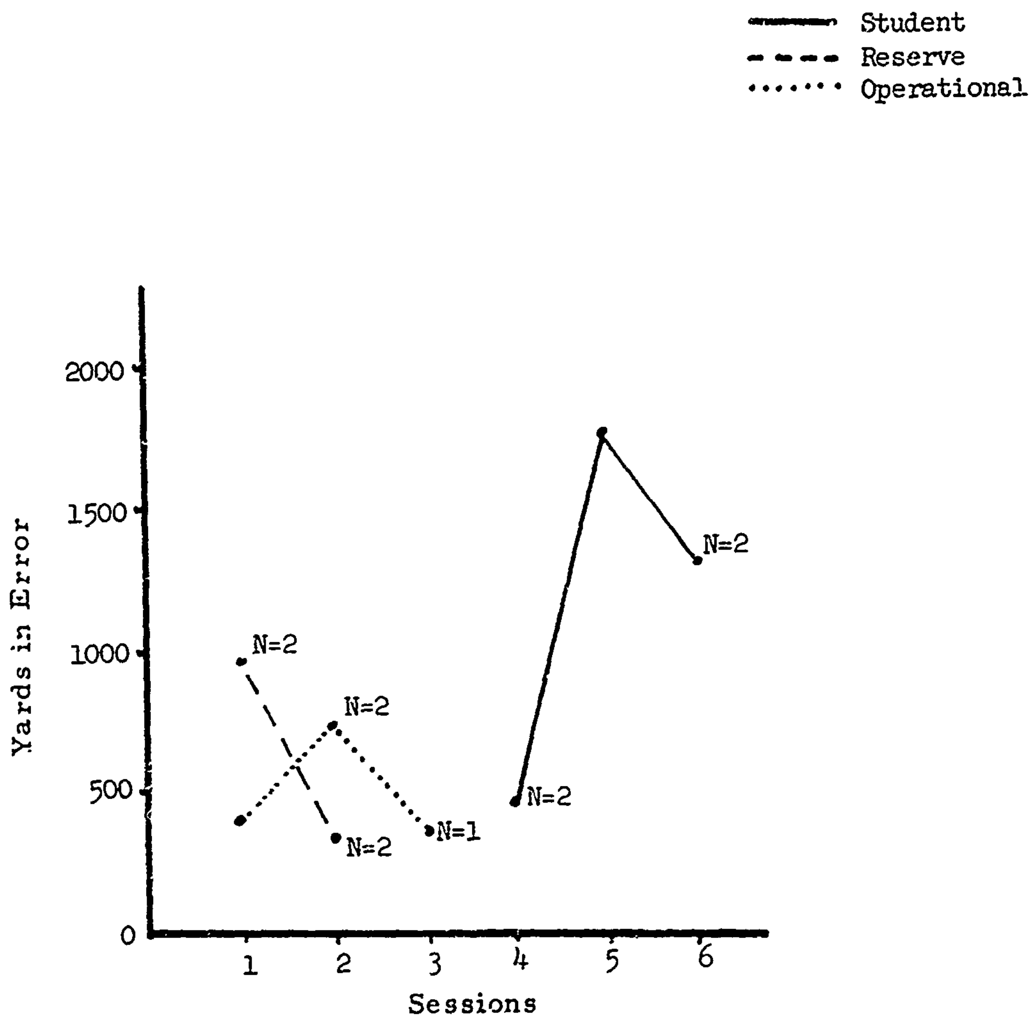


Figure 3. TACCO, Measure B: Buoy Pattern Accuracy During JULIE, Range Error, Population Curves

Measure C: Instructor Ratings of TACCO Performance. Instructors rated all TACCOs at the close of each observed session, on the basis of each TACCO's performance during that session. The scale used was the four point scale normally used by the instructors in evaluating simulator and in-flight performance. This was modified by providing for scaling of operator performance within a given category. After preliminary analysis, the four point scale was reduced to a three point scale because of the absence of any ratings in the fourth (U or unsatisfactory category). (See Instructor Rating Package, Appendix A.)

The subject populations for this measure of TACCO performance are, again, not sufficiently large to draw more than extremely tentative conclusions (see Figure 4). It appears that the most evident change occurred in the student population between the first and fourth sessions; i.e., the trainer was most effective for initial training of student TACCOs.

The interpretation of this subjective measure of TACCO performance was seriously confounded by the fact that TACCOs were not consistently trained or rated by the same instructor. Due to variations in instructor and student schedules (neither controllable by the investigators) this condition prevailed throughout data collection.

Summary of TACCO Performance. The small sample sizes and the incomplete distribution across sessions for the three populations render any conclusions regarding trainer effectiveness for the TACCO position extremely tentative. The data in all cases, even for the curves presented, were quite erratic. Such unreliability, or inconsistency, could be the result of: (1) unidentified factors operating in a complex environment, (2) these data being an accurate description of tolerances within which quantitative operational criteria are met, and/or (3) insufficient and/or inappropriate training for a very demanding position.

Whatever the case, TACCO training requires a further evaluation based on a more adequate data sample if an optimal training program for the TACCO position is to be determined. Such an evaluation should seek information which would relate to at least the three possibilities posed in the preceding paragraph. It may be found that the trainer is effective primarily for initial training, a possibility suggested by the instructor ratings of performance (Figure 4). On the other hand, considering the complex and demanding nature of performance in the TACCO position, it may be that more extensive and/or modified use (e.g., individual training sessions) of the device would increase TACCO performance effectiveness.

NO. 3 OPERATOR.

Measure A: Range Error to Target. This measure is a comparison of the actual range of the target when first detected on radar versus the range as called out by the No. 3 operator. In essence it reflects the operator's error in perceiving and interpreting the range data on his scope.

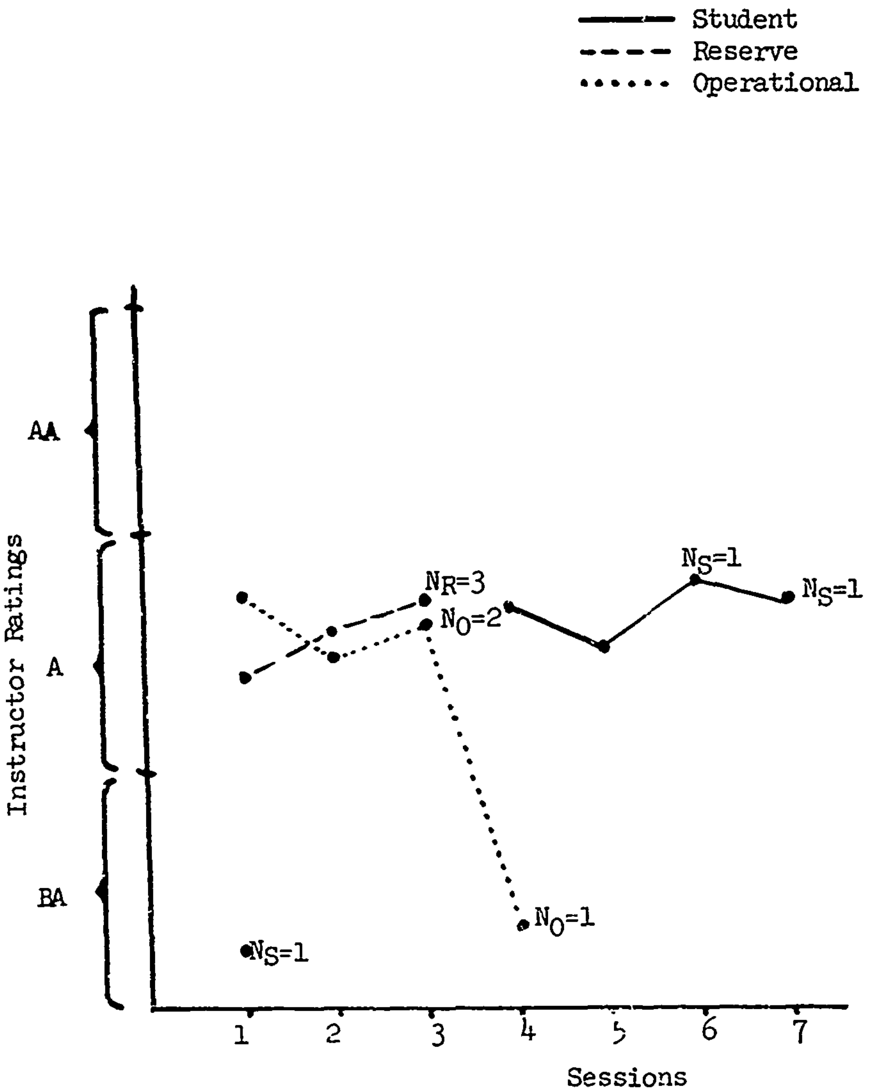


Figure 4. Instructor Ratings of TACCO Performance, Population Curves

The population curves in Figure 5 demonstrate a certain amount of learning. With the exception of that of the Reserve group (which is sufficiently aberrant to warrant discounting) both populations show improved performance as a function of repeated trainer exposure.

Measure B: Time Between Information Updates. This index measures the time between the operator's previous verbal report (bearing/range) and the present report. This measure requires careful interpretation because of the factors which could possibly influence it (discussed in Section I). It should be interpreted only in terms of whether or not substantial changes occurred in training in the operator's time between reports.

The population curves shown in Figure 6 indicate a tendency for time between updates to decrease as a function of training.

Measure C: Target Detection Probability, MAD. This measure indicates the probability that a target will be detected by the operator during MAD operation. The higher the ratio (number of targets detected over opportunities for detection), the better performance is.

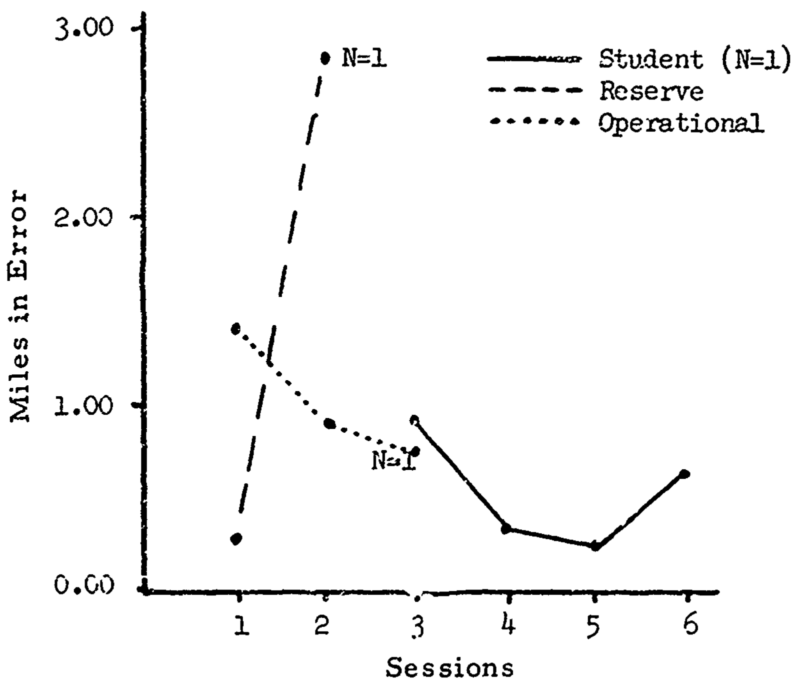


Figure 5. No. 3 Operator, Measure A: Range Error to Target During LOFAR and CODAR, Population Curves

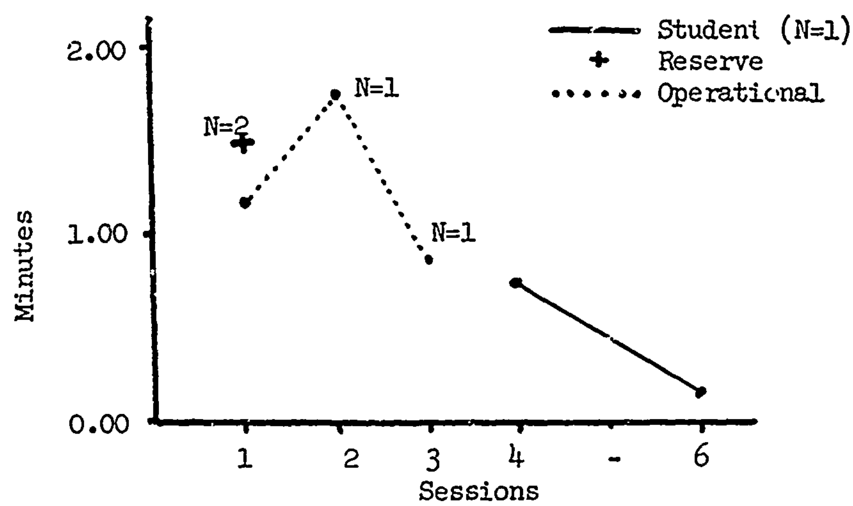


Figure 6. No. 3 Operator, Measure B: Time Between Information Updates, Population Curves

The population curves in Figure 7 indicate that both operational and reserve personnel improve substantially over two training sessions. Student performance in sessions 3 and 4 shows no change. Again, the population sample sizes negate the establishment of any real confidence in these results.

Subjective Measure: Instructor Ratings of No. 3 Operator Performance. The population curves for the instructor ratings shown in Figure 8 depict an improvement over sessions as a function of training. The student population shows slight improvement, while more substantial gains are shown for operational and reserve personnel.

Summary of No. 3 Operator Performance. Evidence has been presented which indicates that the S-2E trainer may be effective for certain No. 3 operator performance criteria. The results shown are considered to be tentative due to the very small sample sizes available for analysis.

NO. 4 OPERATOR.

Measure A: Time to Classify Target, LOFAR. The measure presented in the following figures (Figures 9-1 and 9-2) represents the time between detection of the target, as indicated by a verbal report from the operator, and the time he reports a classification of the target as submarine, nonsubmarine, surface ship, etc.

A definite trend toward improved performance is evident in both the generalized learning curve, Figure 9-1, and in the performance of the operational and student populations (Figure 9-2). No appreciable effect on reserve performance was observed.

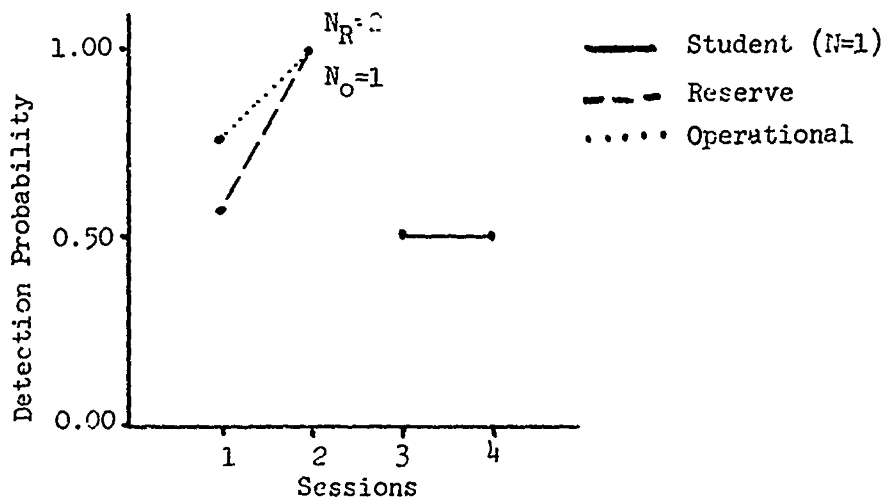


Figure 7. No. 3 Operator, Measure C: Target Detection Probability During MAD, Population Curves

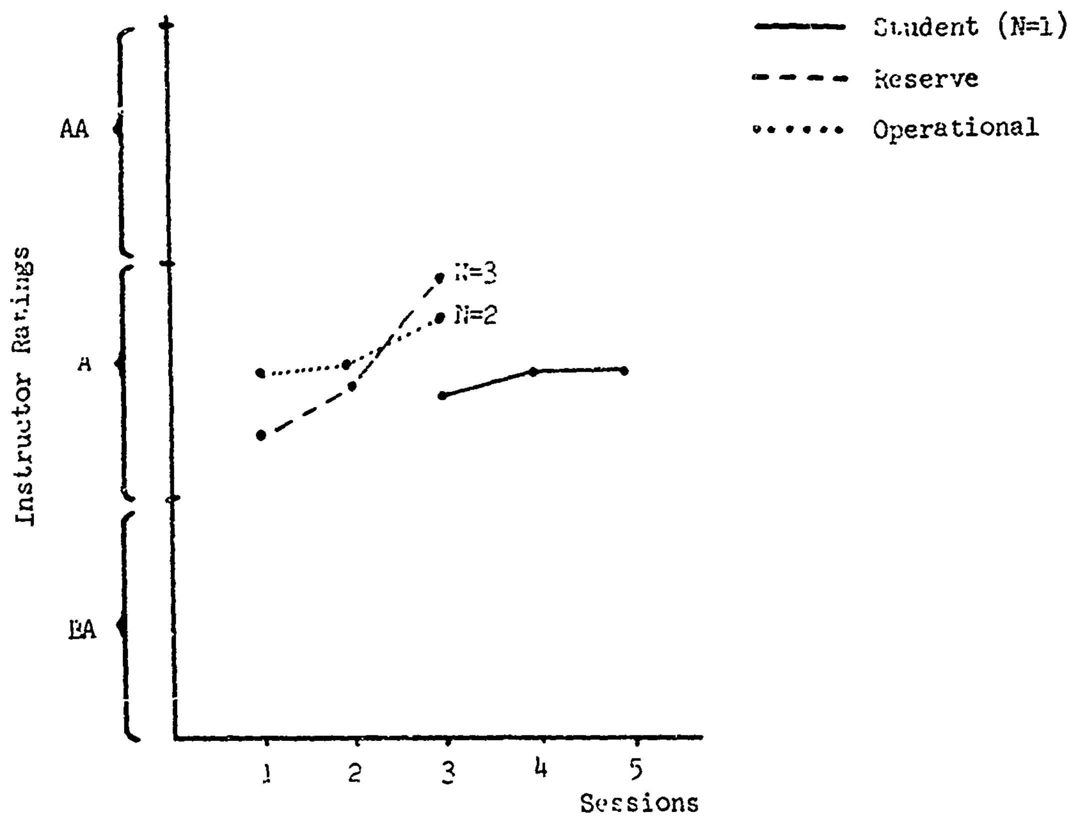


Figure 8. Instructor Ratings of No. 3 Operator Performance, Population Curves

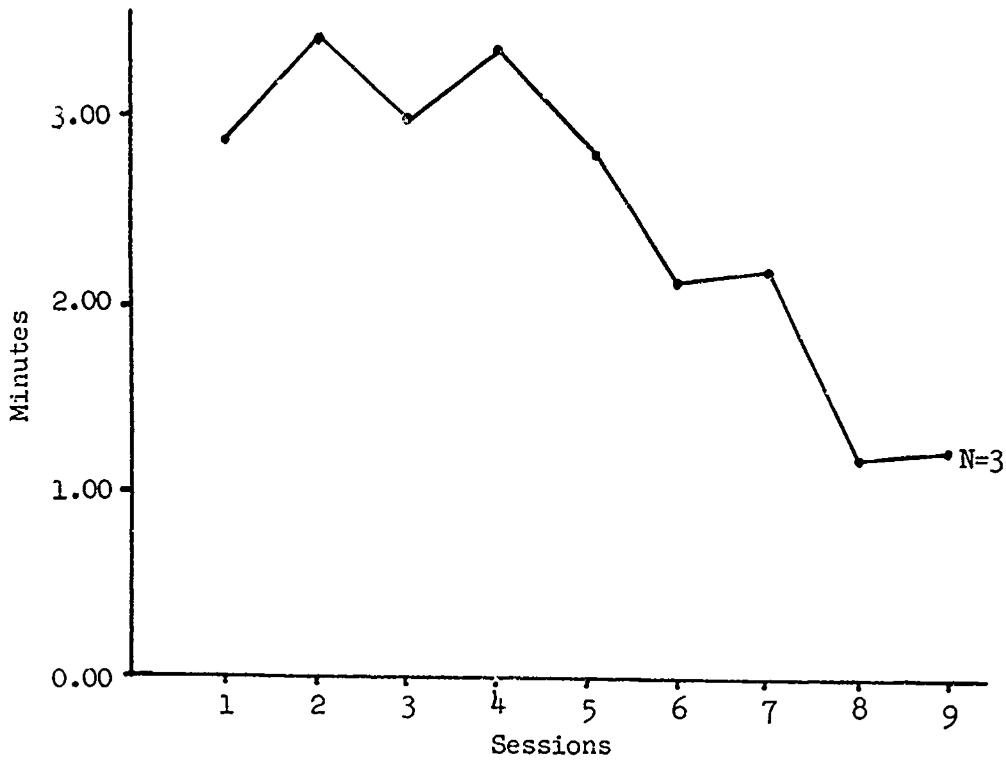


Figure 9. No. 4 Operator, Measure A: Time to Classify Target During LOFAR, Generalized Learning Curve (Part I)

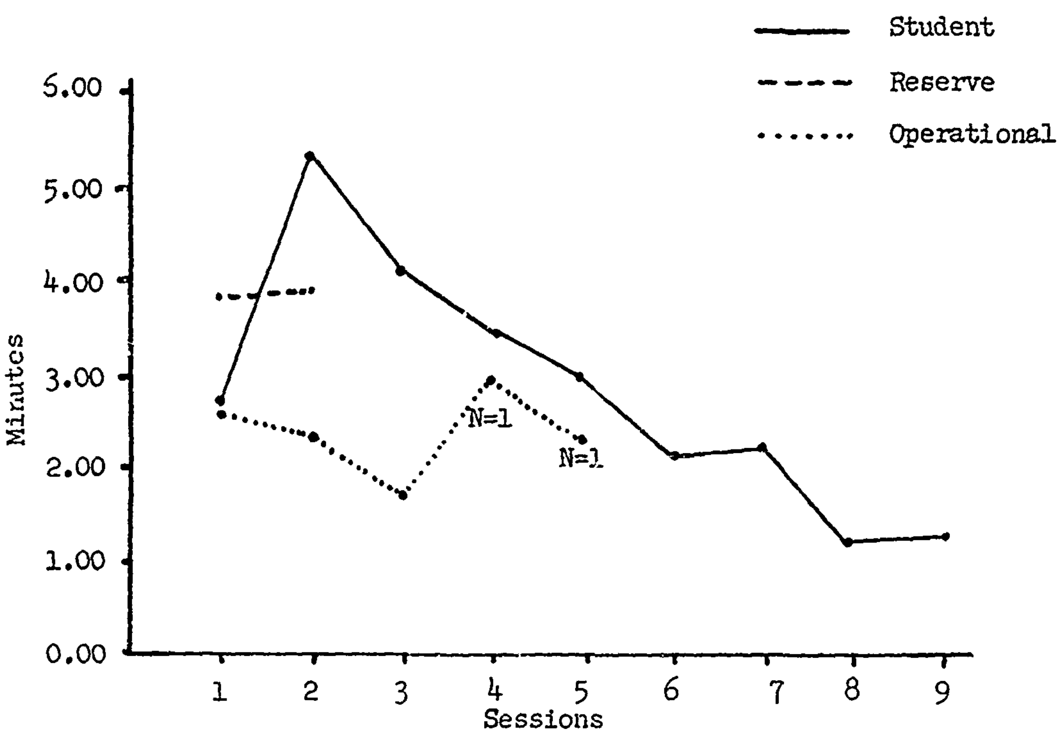


Figure 9. No. 4 Operator, Measure A: Time to Classify Target During LOFAR, Population Curves (Part II)

Measure B: Classification Accuracy, LOFAR. Measure B describes classification accuracy: that is, the accuracy with which operators can categorize the target as submarine or nonsubmarine and identify the type of target as American, Soviet, etc. This measure is a binary one, with "0" being assigned for correct classification and "1" given for incorrect classification.

A high score on this measure therefore represents inaccuracy, a low score, accuracy. Figures 10-1 and 10-2 indicate that both generalized and population curves demonstrate improvement as a function of training. Operational personnel improve more rapidly than students (asymptotic performance reached on session 2 versus session 4). Analysis of the variables impacting upon classification accuracy does not provide any explanation for the degradation in reservist performance between sessions 1 and 2.

Measure C: Range Error to Target, JULIE. This measure is a comparison of the operator's verbal report of the range from the target to a particular buoy (based on his interpretation of his recorder traces) with the computer readout of that range (actual range).

The generalized learning curve (Figure 11-1) indicates marked improvement in performance; this is supported by the individual population curves shown in Figure 11-2. While negative effects are visible at the second and fifth sessions, they are not significant enough to detract from the above conclusion. Analysis of the data with respect to the impact of difficulty variables did not explain the changes in student performance in the fourth and fifth sessions; however, this may be solely a function of the paucity of available data.

Subjective Measure: Instructor Ratings of No. 4 Operator Performance. The generalized curve in Figure 12-1 indicates that there is an overall improvement in No. 4 operator performance as a function of training (ratings rise from Average to a point midway between Average and Above-Average by the ninth session). This trend in rated performance is substantiated by the curves for student and reserve populations (Figure 12-2). While the ratings for the operational population indicate some loss, the latter sessions of operational performance deal only with one to three individuals. Based on the instructor ratings, it seems reasonable to state that No. 4 operators, particularly students, made progress as a function of training on this device. These ratings confirm the results of the objective measures.

Summary of No. 4 Operator Learning. Examination of the various curves for the No. 4 operator indicates that on the whole considerable learning has occurred, primarily for students, but also for other population samples. The availability of a larger sample of No. 4 operators in all populations allowed the possibility of confounded or artifactual

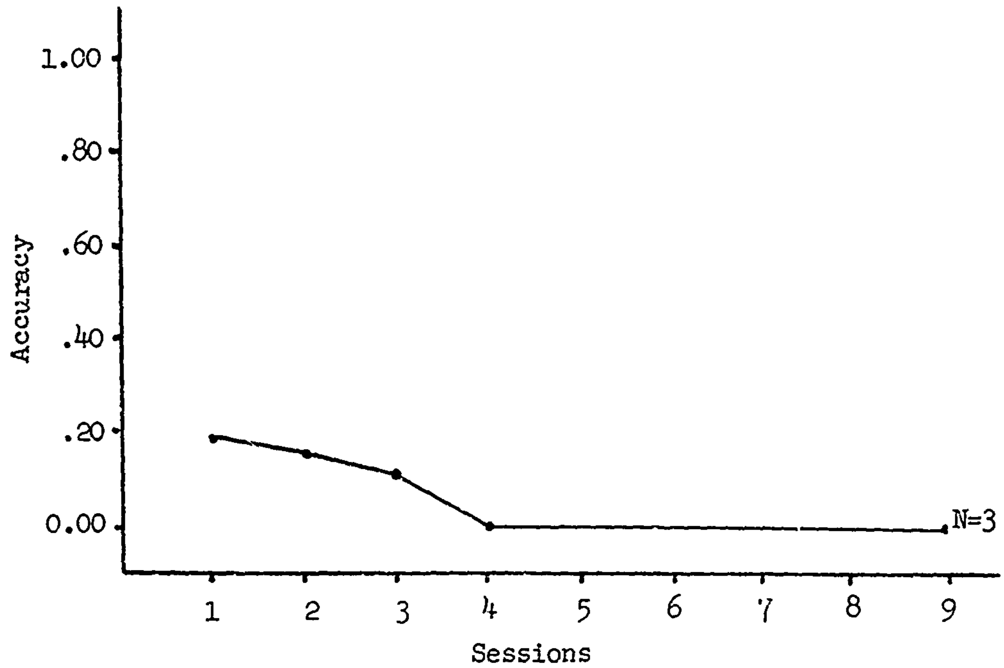


Figure 10. No. 4 Operator, Measure B: Classification Accuracy During LOFAR, Generalized Learning Curve (Part I)

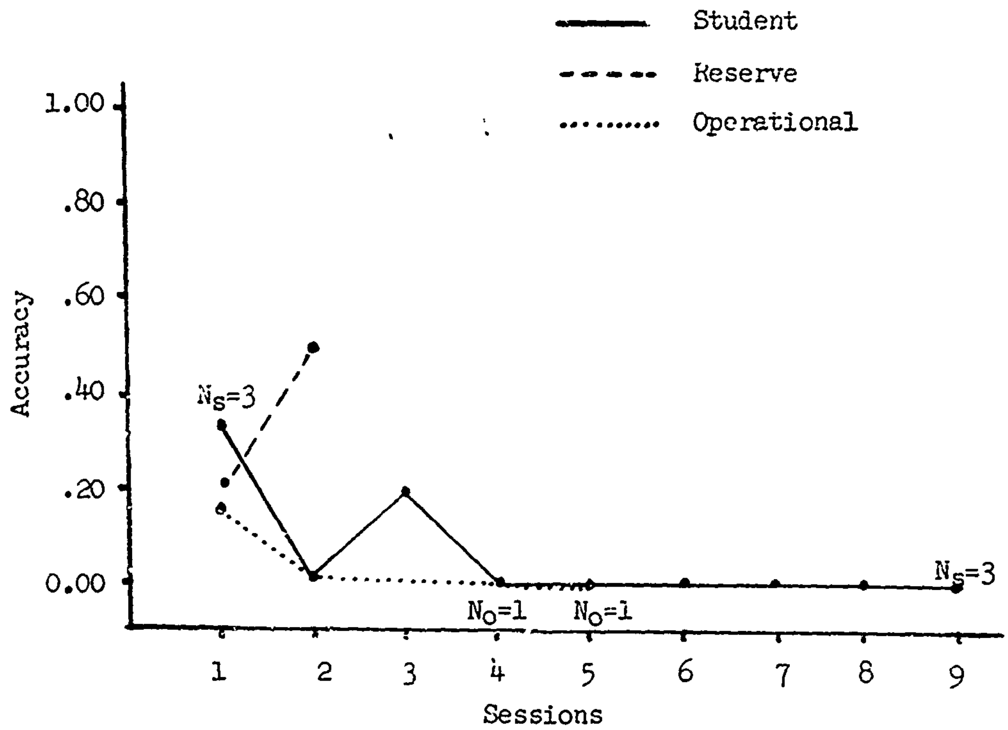


Figure 10. No. 4 Operator, Measure B: Classification Accuracy During LOFAR, Population Curves (Part II)

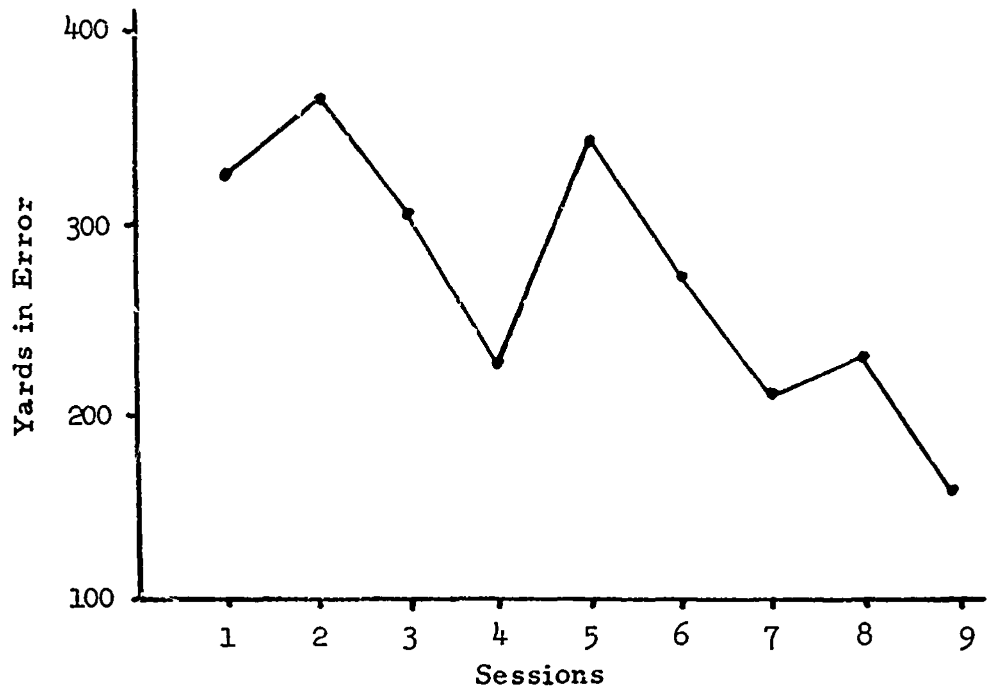


Figure 11. No. 4 Operator, Measure C: Range Error to Target During JULIE, Generalized Learning Curve (Part I)

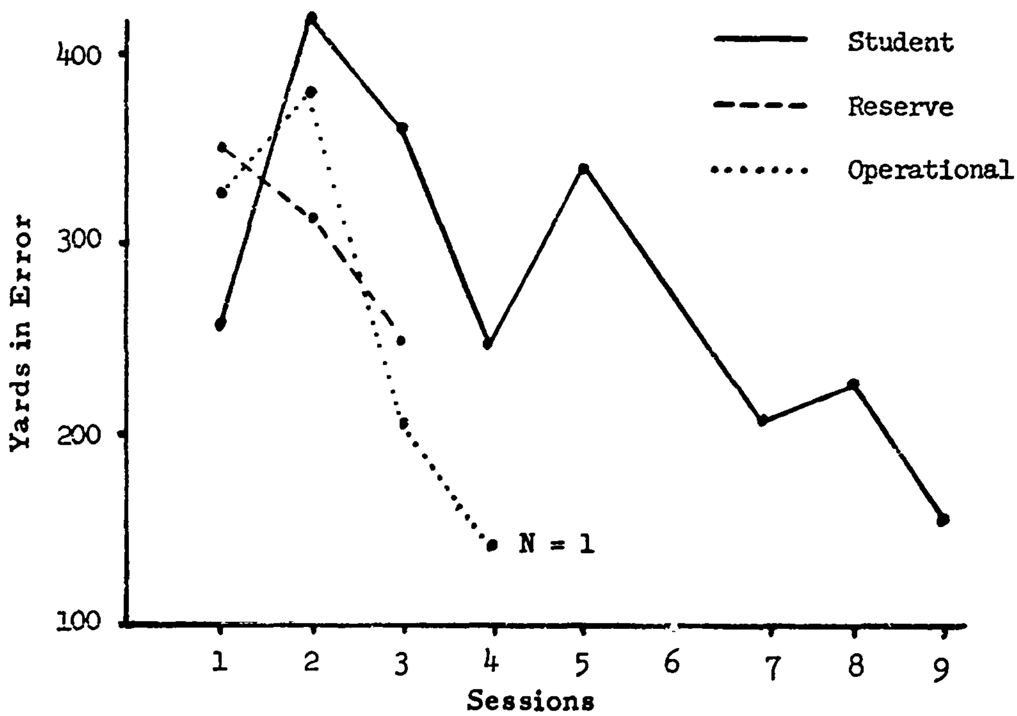


Figure 11. No. 4 Operator, Measure C: Range Error to Target During JULIE, Population Curves (Part II)

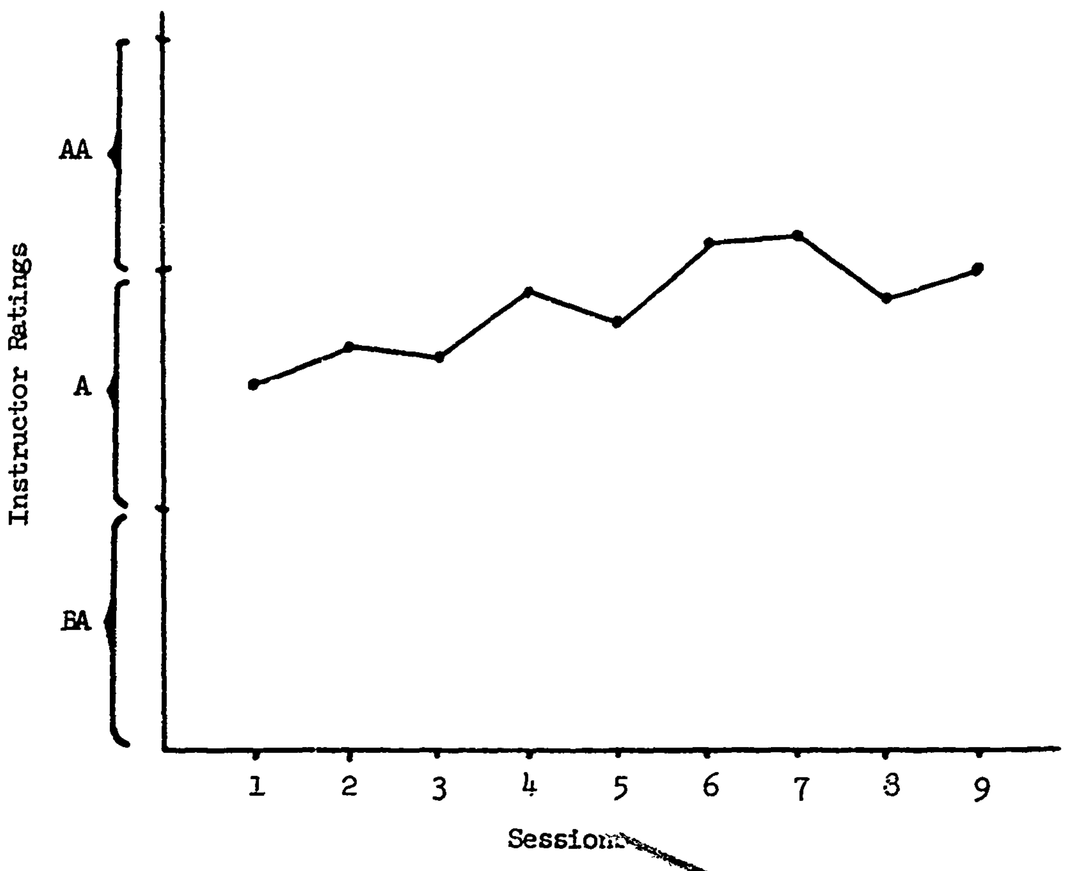


Figure 12. Instructor Ratings of No. 4 Operator Performance, Generalized Learning Curve (Part I)

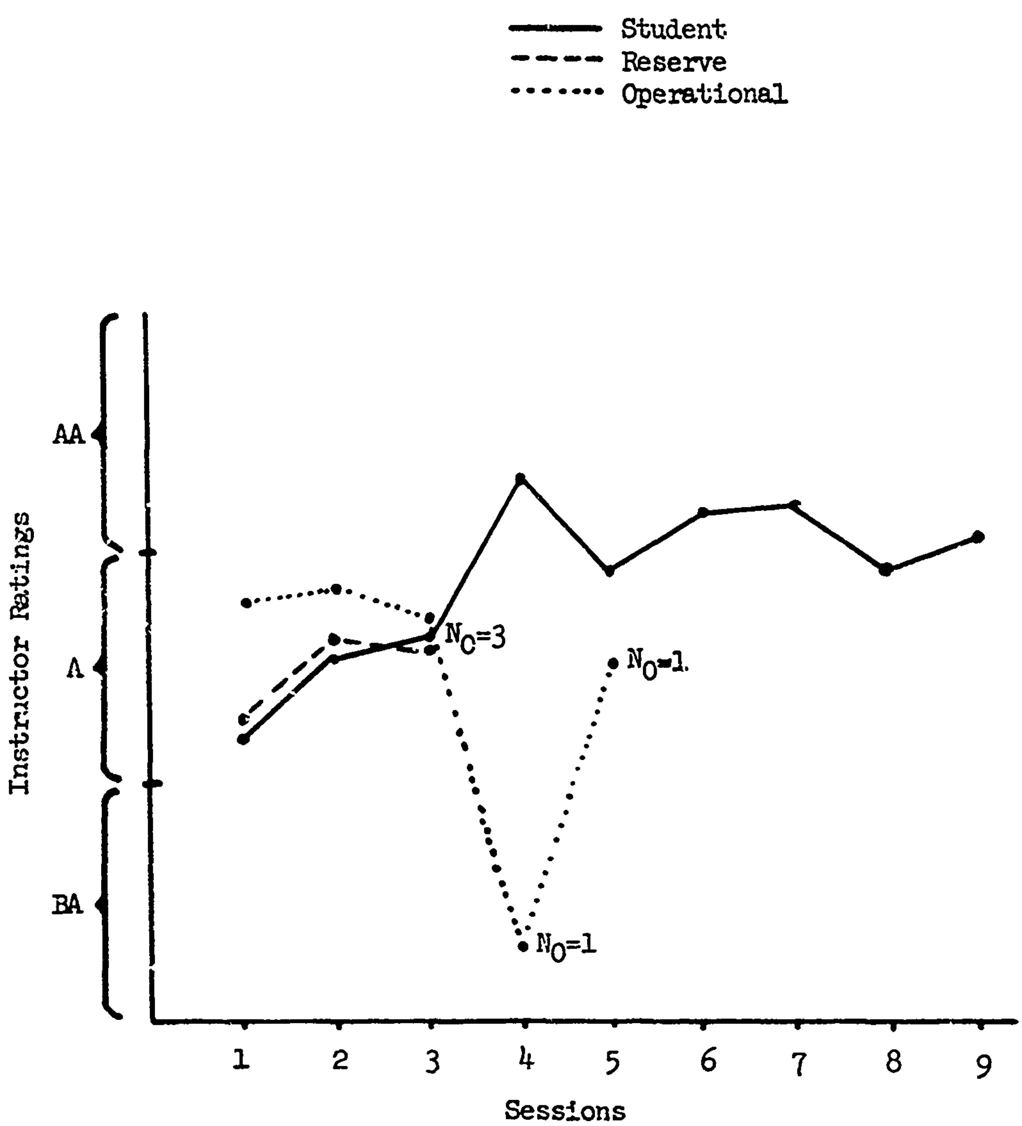


Figure 12. Instructor Ratings of No. 4 Operator Performance, Population Curves (Part II)

learning effects to be analyzed far more extensively than was possible for other positions (see pp 34-36). As a result, a degree of confidence can be placed in the interpretation of the presented curves as demonstrations of trainer effectiveness. The other measures of No. 4 operator performance listed in Table 2, but not presented here, were either comparatively insensitive to trainer use or indicated highly erratic performance with an insufficient sample size to accurately determine why.

HOW DOES LEARNING ON THE DEVICE COMPARE AMONG STUDENTS, OPERATIONAL CREWS, AND RESERVISTS?

Although this question has been answered by the population curves in the previous section, it will be helpful to summarize what has been learned. The three populations can be considered in two different ways: by position (TACCO, No. 3 operator, No. 4 operator) or by experience or background (students, operational personnel, reserves). A review of the individual learning curves suggests very strongly that the No. 4 operator position receives maximum training benefit from the training device, while the No. 3 operator and the TACCO receive less. The demonstration of trainer effectiveness for the TACCO and No. 3 operator may, of course, be prejudiced by the limited (small sample size) and restricted (fewer sessions sampled) data available for analysis.

In summary, it can be seen that for the operational and student populations, training on Device 2F66A resulted in improved performance. The reservists did not manifest as much improvement, but this may be attributable to either the "hump" phenomena observed in the other populations, or it may be the fatigue effect brought on by the active annual training schedule. In any event, the observation span for reservists is too small to warrant any explicit conclusion.

Although the utility of trainers for student personnel is usually granted, the utility of trainers for operational personnel is sometimes contested. In line with this, it should be pointed out that the range between best and poorest operational personnel performance was very large for most of the performance measures. This, coupled with the improvement indicated by certain learning curves, would seem to argue for more systematic trainer use on the part of operational personnel.

HOW IS LEARNING INFLUENCED BY VARIOUS TRAINING INPUTS?

A number of inputs to the training situation (described in Section I) may influence the learning produced by the training device. In the case of Device 2F66A, these include variations in such inputs as number of targets, target maneuvers, etc., which are expected to alter the difficulty level of the mission. It would be expected, as a result of their differing functions and tasks, that different crew members would be affected differently and to different extents as a function of the combination of training inputs utilized.

Tables 6, 7, and 8 describe the training inputs to each training session for each population. The values presented are the number of teams (or individuals, where no other positions were occupied) constituting the sample population per session, the mean frequency of inputs per given session, and the range of input frequencies across the session sample.

In the tables mentioned, it is immediately apparent that there was considerable variation in the combination and extent of training inputs to each of the teams constituting the individual training session samples. The variation is substantial both within sessions and between successive sessions. The variation in difficulty level within and between sessions arose because instructors characteristically tailor the individual training session to what they conceive to be the needs of the particular trainee or crew.

The variability of the inputs between and within sessions suggests the possibility that instructor judgments about training requirements for the individual student may have been, if not inaccurate, at least random. One would expect a more regular pattern of training input usage and a tendency toward increased usage as training progresses—particularly for the students. Examination of cell means across sessions, and ranges within cells in Tables 6, 7, and 8, indicate that neither tendency appears on a regular basis. On the contrary, the difficulty of successful performance on the missions, as affected by training inputs, appears extremely variable. One might also question the advisability of the apparent infrequent use of the contingency variables (recorder failure, sonobuoy failure, receiver malfunction, loss of contact) for the operational and reserve populations. Effectiveness in the operational environment, i.e., mission success, may depend on adequate performance in the face of these contingencies. This implies a need to deal effectively with a degraded system if it becomes necessary.

The performance data was examined with respect to the training inputs to determine whether effects on performance could be identified. The results of the analyses suggest that the effect (not necessarily negative) of training inputs on subject performance is a complex multivariate relationship possibly including the following variables: point in training (i.e., both population and session), idiosyncratic characteristics of the crew member, and the particular position measure under discussion. In no case was a consistent relationship found which changed the implications of the curves in the previous section. It is suggested that this issue be examined more closely in future studies. In the present case, considering the almost random nature of the inputs across the data sample, the small sample size available for analysis, and the lack of consistent relationships, a final answer describing the effects of trainer inputs on performance can not be given.

TABLE 6. FREQUENCY OF TRAINING INPUTS OVER SESSIONS FOR THE STUDENT POPULATION LOCATED IN SAN DIEGO

	Student Population								
	Sessions, Session Sample Size								
Trainer Inputs	1 N = 5	2 N = 4	3 N = 7	4 N = 9	5 N = 7	6 N = 8	7 N = 7	8 N = 5	9 N = 2
Number of targets	1.4* 1-2	1.0 1**	1.3 1-3	1.4 1-3	1.1 1-2	1.8 1-3	1.0 1	1.8 1-3	1.0 1
Target speed changes	0.4 0-1	2.5 2-3	0.6 0-3	0.6 0-2	0.6 0-2	0.8 0-2	1.0 0-2	1.2 0-3	2.0 2
Target course changes	0.0 0	1.0 0-2	1.0 0-3	0.6 0-3	0.6 0-1	1.3 0-3	1.1 0-3	1.2 0-2	0.5 0-1
Target depth changes	0.0 0	0.0 0	0.0 0	0.3 0-2	0.4 0-3	0.0 0	0.0 0	0.0 0	0.0 0
Recorder failure induced	0.0 0	0.5 0-2	0.3 0-1	0.1 0-1	0.0 0	0.0 0	0.3 0-1	0.0 0	0.0 0
Sonobuoy failure induced	0.0 0	0.0 0	0.1 0-1	0.1 0-1	0.0 0	0.0 0	0.0 0	0.0 0	0.0
Receiver malfunction induced	0.0 0	0.0 0	0.1 0-1	0.0 0	0.0 0	0.0 0	0.1 0-1	0.0 0	0.0 0
Loss of contact induced	0.2 0-1	0.0 0	0.1 0-1	0.0 0	0.1 0-1	0.0 0	0.0 0	0.0 0	0.0 0

*Legend: Mean frequency of input across session sample
Range of inputs across session sample

**A single range value indicates the frequency of input was constant across the sample for that session

TABLE 7. FREQUENCY OF TRAINING INPUTS OVER SESSIONS FOR THE RESERVE POPULATION LOCATED IN SAN DIEGO

	Reserve Population		
	Sessions, Session Sample Size		
Trainer Inputs	1 N = 14	2 N = 10	3 N = 4
Number of targets	1.0* 1**	1.0 1	1.0 1
Target speed changes	0.8 0-5	0.9 0-4	0.0 0
Target course changes	1.7 0-6	1.6 0-5	1.3 0-2
Target depth changes	0.6 0-2	0.2 0-1	0.5 0-1
Recorder failure induced	0.0 0	0.0 0	0.0 0
Sonobuoy failure induced	0.2 0-1	0.3 0-1	0.0 0
Receiver malfunction induced	0.0 0	0.0 0	0.0 0
Loss of contact induced	0.1 0-1	0.0 0	0.3 0-1

*Legend: Mean frequency of input across session sample
Range of inputs across session sample

**A single range value indicates the frequency of input was constant across the sample for that session

TABLE 8. FREQUENCY OF TRAINING INPUTS OVER SESSIONS FOR THE OPERATIONAL POPULATION LOCATED IN SAN DIEGO

	Operational Population				
	Sessions, Session Sample Size				
Trainer Inputs	1 N = 32	2 N = 16	3 N = 6	4 N = 1	5 N = 1
Number of targets	1.3* 1-3	1.1 1-3	1.2 1-2	1.0 1**	1.0 1
Target speed changes	0.8 0-3	0.7 0-2	0.7 0-2	1.0 1	0.0 0
Target course changes	1.0 0-5	1.8 0-7	2.3 0-5	1.0 1	0.0 0
Target depth changes	0.3 0-1	0.0 0	0.3 0-2	0.0 0	0.0 0
Recorder failure induced	0.0 0**	0.0 0	0.0 0	0.0 0	0.0 0
Sonobuoy failure induced	0.0 0	0.0 0	0.2 0-1	0.0 0	0.0 0
Receiver malfunction induced	0.0 0	0.0 0	0.0 0	0.0 0	0.0 0
Loss of contact induced	0.0 0	0.0 0	0.2 0-1	0.0 0	0.0 0

*Legend: Mean frequency of input across session sample
Range of inputs across session sample

**A single range value indicates the frequency of input was constant across the sample for that session

HOW IS LEARNING INFLUENCED BY THE INTRODUCTION OF TEAM TRAINING AND THE SEQUENCING OF SIMULATOR SESSIONS WITH FLIGHT TRAINING SESSIONS?

HOW IS LEARNING INFLUENCED BY THE INTRODUCTION OF TEAM TRAINING? Device 2F66A is used both as an individual and as a team trainer. In the individual training situation, the only crew member present might be the No. 4 operator. The instructor then provides all TACCO/pilot and mission inputs to the operator, so that a particular phase can be replicated. In a team session two or more members of the crew are present and perform as interactive members of that crew. It is important to remember that in both situations (individual or team) an operational mission phase is simulated; however, the degree of operational realism and particularly the coordination required of the individual crewman is greater in the team setting.

On the whole, except for TACCO students, the trainer was not utilized as a team trainer for the student population at San Diego, although it was used primarily in this mode by the other two populations. As can be seen by reference to Figure 13, only two of the nine possible sessions at most were team sessions for only five of the student No. 4 operators. Further, as a result of scheduling, in no instance was the trainer used as a full team trainer during any of the sessions observed; i.e., the No. 3 position was never filled while the other positions trained, nor did the only student No. 3 operator ever perform as a member of a team In Key West, students did not perform as part of a team at any time during the observation period.

With the exception of one measure, there was no consistent deviation in performance level across the student No. 4 operator population to the introduction of team training. It appeared that student performance was disturbed primarily during the first half of a team training session (i.e., measures of LOFAR and CODAR performance), but that the direction (positive or negative) and amount of response were highly idiosyncratic. Team training effects on performance during the latter half of the training session were not at all obvious. The task performance most obviously and consistently affected by the introduction of team training was that described by Measure 1 (listed in Table 2): "Time to Detect Target During LOFAR."

In Figure 13, those sessions in which the student No. 4 operator performed as a member of a team have been indicated by T_1 (for the first team training session), and T_2 (for the second team training session). The curves describe the performance of individual No. 4 students on Measure 1. It can be noted that student performance was degraded in four out of five T_1 sessions, while in the fifth T_1 session an unreasonably quick response was made (student F)—unreasonable in the sense that a decision regarding signal to noise probabilities could

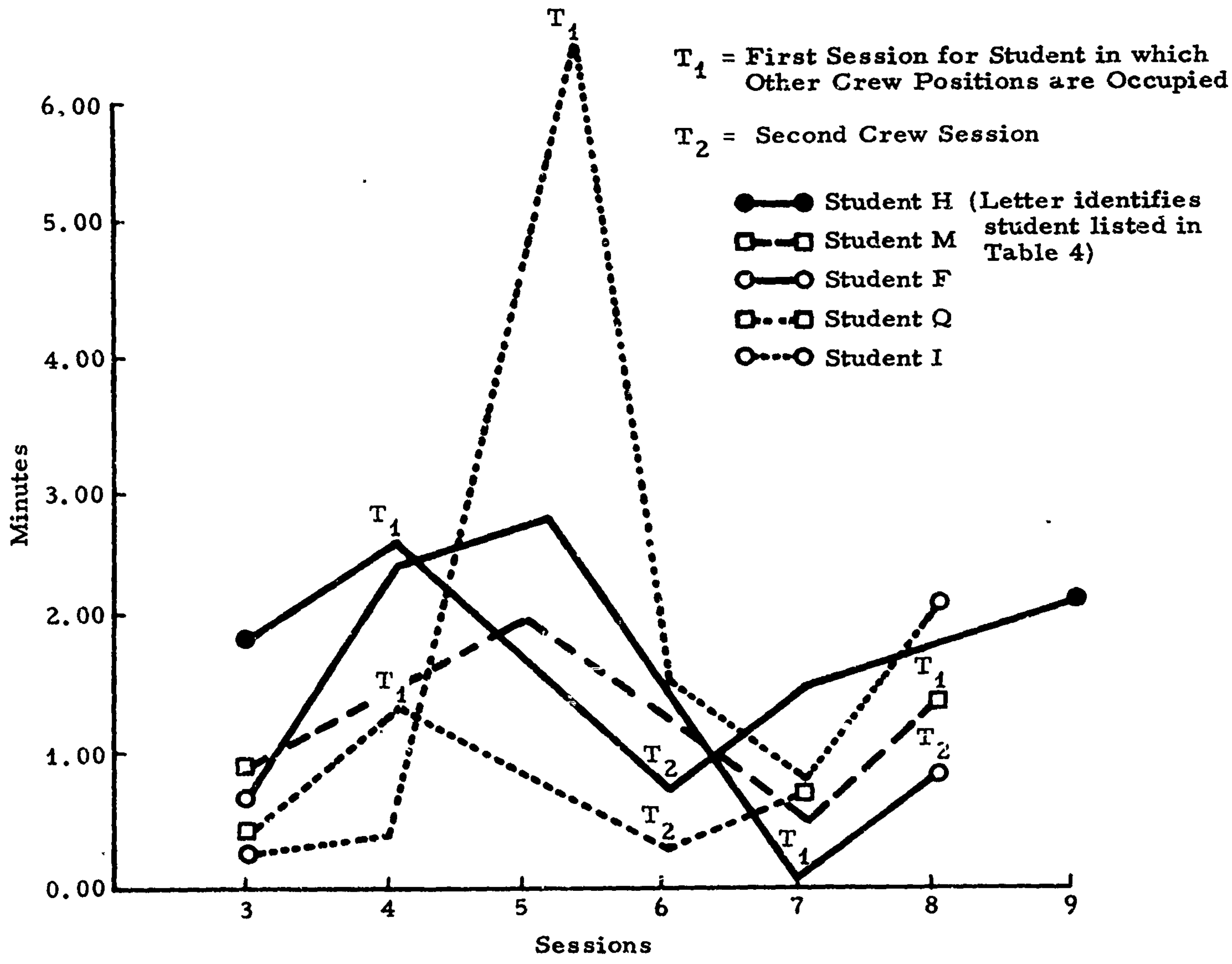


Figure 13. Effect of Initial Crew Training Sessions on Student No. 4 Operator Performance During LOFAR as Described by Measure 1: Time to Detect Target

not have been made. On the second team training session, two of the three students improved their performance; the third student (student F again) required longer to respond on the T_2 session than on the T_1 session, but as indicated previously, the performance was at a more reasonable level. Assuming that the correlation of events (T_1, T_2) and changed performance levels is not a random one, then one can say that the use of the trainer as a team trainer is effective for this measure in that the student appears to adjust to the presence of team members. The first time the No. 4 student operator performed with a team, he performed his task less adequately, but on the second team session, he performed at a level more closely associated with his pre-team performance.

COMMUNICATION AS A REFLECTION OF TEAM TRAINING. Since one of the explicit goals of WST training is to improve crew "coordination," the investigators decided that some measure of coordination which could demonstrate the effect of the training device on this aspect of team performance was required.

Since crew interaction in the S-2E aircraft was accomplished almost exclusively via verbal communications, the decision was made to explore changes in the relative frequency of communications categories that occur in training. Even if one does not regard communications as an index of crew coordination, changes in communications patterns resulting from training can be viewed in a more specific sense as training in communications procedures. Regardless, therefore, of one's interpretation of the significance of this measure, it is fundamental to adequate performance of the AASW mission.

Two types of communications analyses were performed for each position in the student population. Changes were determined in the relative frequency of the categories utilized in communications across the training sessions (Figures 14 and 15). The categories to which the crew member statements were assigned can be summarized as follows:

a. C1, requests information, opinion, or course of action

b. C2, provides requested information, opinion, or course of action

c. C3, declarative information (including direct commands)

d. C4, verifies information (feedback)

e. C5, acknowledgments

f. C6, irrelevant remarks

-------- MEAN PERFORMANCE, OPERATIONAL NO. 4 OPERATORS

———— SESSION PERFORMANCE, STUDENT NO. 4 OPERATORS

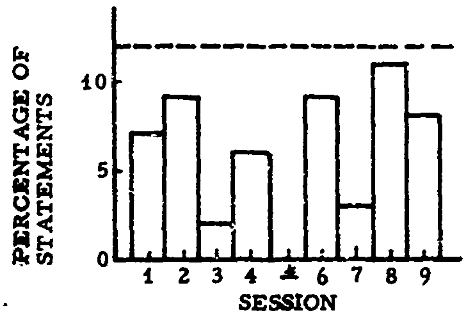


Figure 14-1. C1 Communications

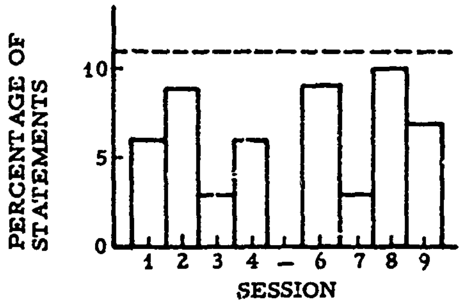


Figure 14-2. C2 Communications

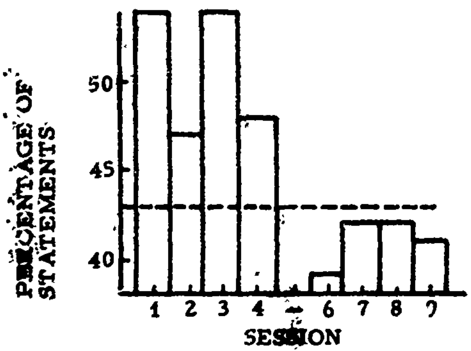


Figure 14-3. C3 Communications

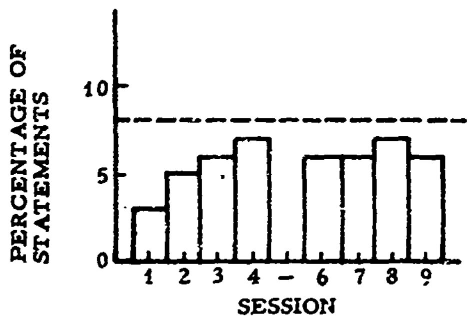


Figure 14-4. C4 Communications

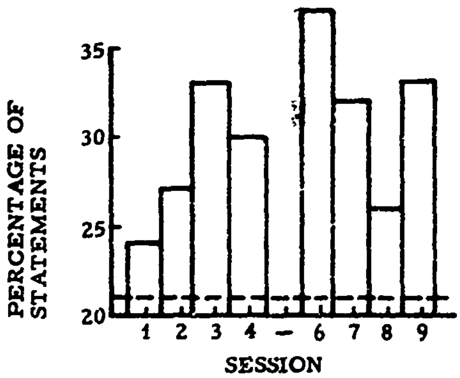


Figure 14-5. C5 Communications

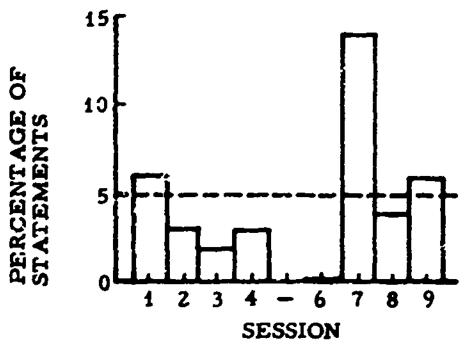


Figure 14-6. C6 Communications

*NOTE: SESSION 5 NOT REPRESENTED DUE TO INSUFFICIENT DATA

**TOTAL PERCENTAGE PER SESSION ACROSS FIGURES 14-1 TO 14-6 EQUALS 100%

Figure 14. Communications Involving Student No. 4 Operators Are Identified by Type (C1-C6) and Compared to the Mean for Operational No. 4 Operators

------ MEAN PERFORMANCE, OPERATIONAL TACCO/ PILOTS

——— SESSION PERFORMANCE, STUDENT TACCO/PILOTS

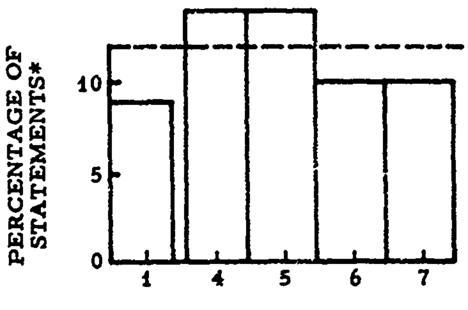


Figure 15-1. C1 Communications

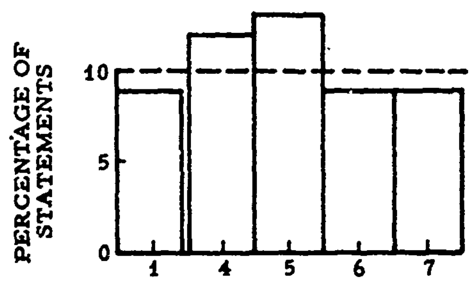


Figure 15-2. C2 Communications

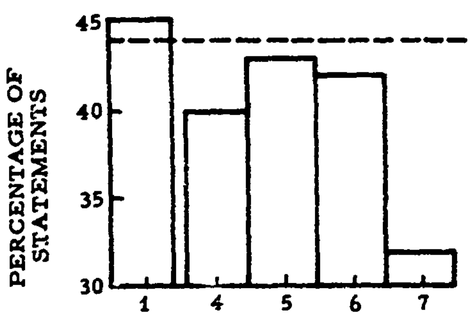


Figure 15-3. C3 Communications

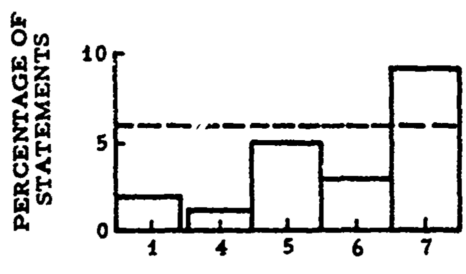


Figure 15-4. C4 Communications

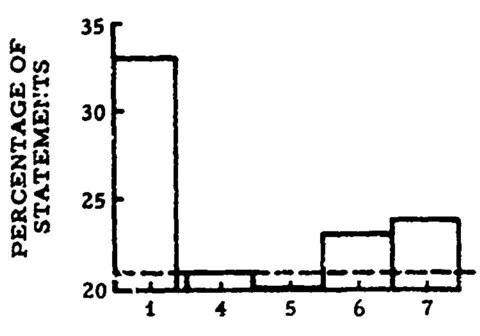


Figure 15-5. C5 Communications

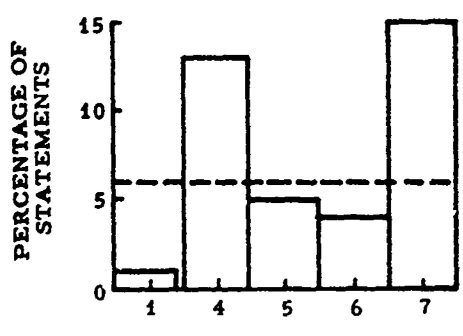


Figure 15-6. C6 Communications

*TOTAL PERCENTAGE PER SESSION ACROSS FIGURES 15-1 THROUGH 15-6 EQUALS 100%

Figure 15. Communications Involving Student TACCO/Pilots Are Identified by Type (C1-C6) and Compared to the Mean for Operational TACCO/Pilot Personnel

The relative frequency with which each category was utilized within a session sample has been expressed as a percentage, where the total percentage summed across the six categories equals 100 percent.

Also, changes were determined in relative frequency of communication direction over training sessions where a statement is described as a directional exchange between two positions (Figures 16 and 17). For example, the descriptor (No. 4 operator →instructor) would describe a statement made by the No. 4 operator to the instructor. As noted before, communications identified as "TACCO/pilot" are communications concerned with the TACCO function. The relative frequency is again expressed as a percentage, where the total percentage summed across communication exchanges involving a particular position equals 100 percent.

In all cases the session values determined for the student population were compared with the average percentage values found for the operational population sessions.

Changes in Relative Frequency of Student use of Communications Categories During Training. Figure 14 presents histograms describing the percentage of statements by communications category of student No. 4 operators across the nine training sessions (session 5, however, was not included because there was only one observation of the fifth training session for student No. 4 operators). The use of histograms, rather than graphs, is not intended to imply a lack of training continuity across sessions. Rather, the histogram format was selected because the informational content of the data was felt to be more clearly presented in this mode. The percentages represent all communications in which the No. 4 operator was involved, e.g., a communication from the TACCO to the operator and vice versa.

Chi-square analysis of the frequencies represented by these percentages was significant at the 0.01 level ($X^2 = 90.79$, 48 df). This indicates that statistically significant changes in communication patterns occurred during training. Examination of the change in category usage across sessions as shown in Figures 14-1 through 14-6 leads to the conclusion that training may systematically affect two types of student No. 4 operator communications: C3, Declarative Information and C4, Verifies Information. It appears that voluntary inputs of new information or opinion, i.e., C3 communications, gradually decreases and approaches the operational mean for No. 4 operators. Although a lesser change occurs for C4 across sessions, the change in this type of communication is again relatively consistent and approaches the operational mean. No learning trends are obvious for categories C1, C2, C5, or C6. Since most of these training sessions were of the individual practice type (No. 4 student operator working with an instructor simulating other positions), it appears that even in the individual training context one effect of training device use is to influence the pattern of operator communications with other crew positions.

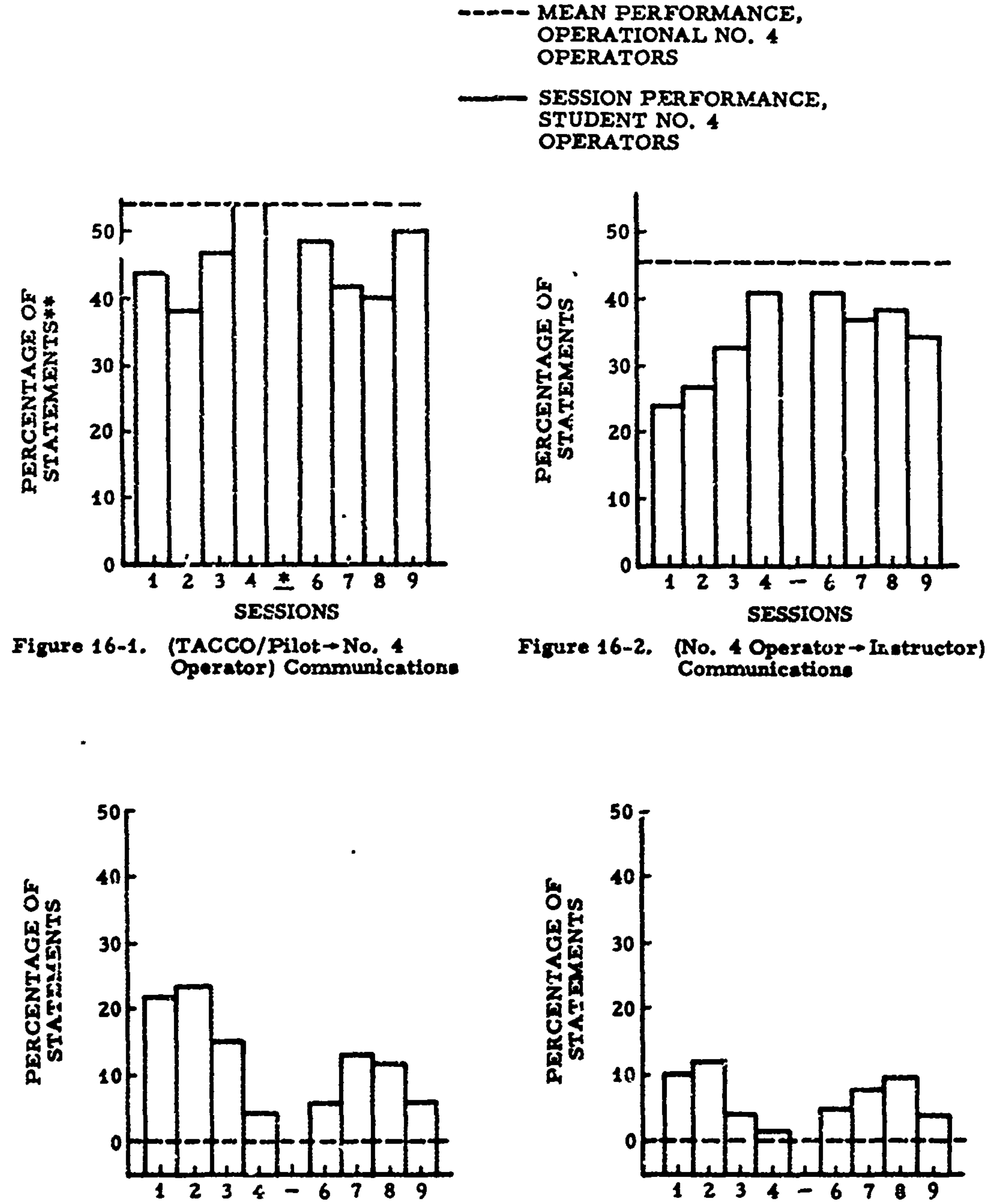


Figure 16-1. (TACCO/Pilot→No. 4 Operator) Communications

Figure 16-2. (No. 4 Operator→Instructor) Communications

Figure 16-3. (Instructor→No. 4 Operator) Communications

Figure 16-4. (No. 4 Operator→Instructer) Communications

*NOTE SESSION 5 NOT REPRESENTED DUE TO INSUFFICIENT DATA

**TOTAL PERCENTAGE PER SESSION ACROSS FIGURES 16-1 THROUGH 16-4 EQUALS 100%

Figure 16. Interpersonal Communications Involving No. 4 Operator Students Across Training Sessions as Compared to the Mean for Operational No. 4 Operators

---- MEAN PERFORMANCE
——SESSION PERFORMANCE

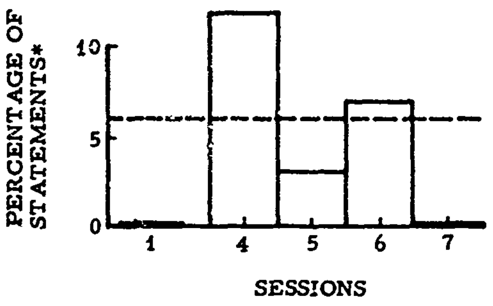


Figure 17-1. (TACCO/Pilot → No. 3 Operator) Communications

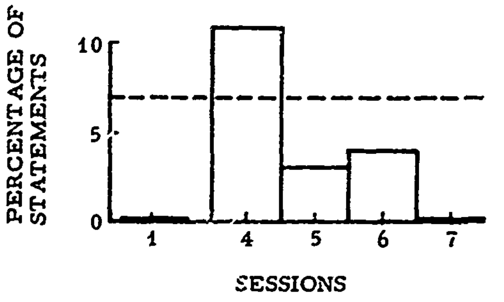


Figure 17-2. (No. 3 Operator → TACCO/ Pilot) Communications

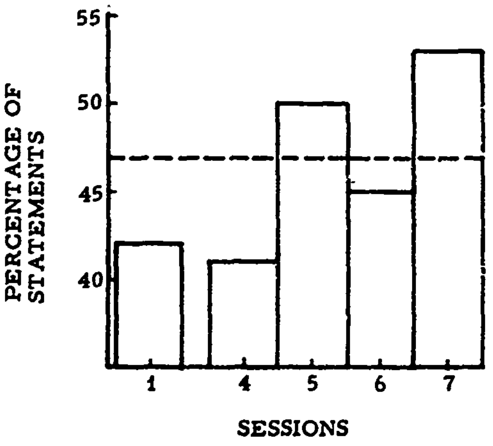


Figure 17-3. (TACCO/Pilot → No. 4 Operator) Communications

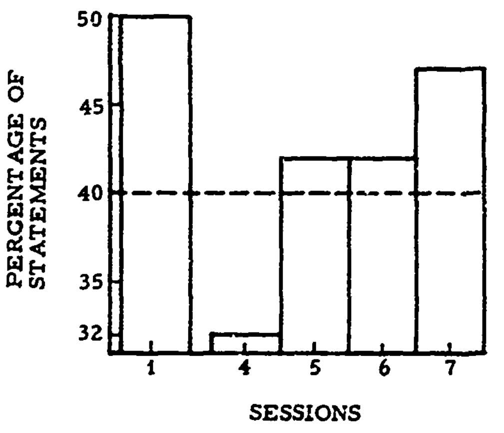


Figure 17-4. (No. 4 Operator → TACCO/ Pilot) Communications

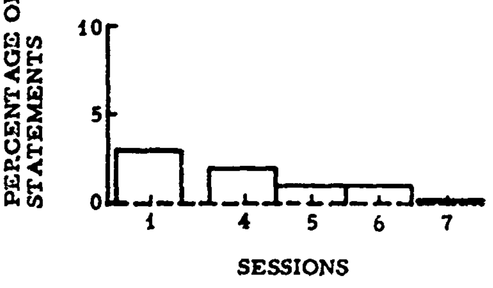


Figure 17-5. (TACCO/Pilot → Instructor) Communications

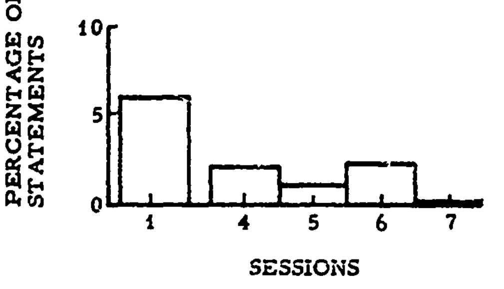


Figure 17-6. (Instructor → TACCO/ Pilot) Communications

*TOTAL PERCENTAGE PER SESSION ACROSS FIGURES 17-1 TO 17-6 EQUALS 100%

Figure 17. Interpersonal Communications Between TACCO and Pilot Positions and the No. 3 and No. 4 Operators, Mean Performance Compared to Session Performance

The overall deviation of student C5 communications from the mean of operational No. 4 operators reflects the observed tendency on the part of the student operator to be aggressively "Naval" in his responses, by acknowledging every command given him by the TACCO. The need for making formal acknowledgments is apparently lessened when members of a crew are familiar with each other's manner of operation.

Figure 15 shows the change in TACCO/pilot student communications during the five training sessions for which data are available. The values represent communication category use in all communications involving interaction between the TACCO/pilot positions and all other positions. The changes represented are quite marked, some of them being consistent and in the direction of meeting the mean of operational personnel, while others are not. As observed for TACCO data presented in previous paragraphs, the observation sample is not sufficiently large to draw more than very tentative conclusions. It appears that the use of C4, Verifies Information, and C5, Acknowledgments (as shown in Figures 15-1 through 15-6) that the change is both somewhat consistent and in the direction of the mean for operational TACCO/pilot personnel. In other words, performance trends are apparent and are in the desired direction. C1 and C2, on the other hand, appear to constitute a relatively constant portion of TACCO/pilot communications, while C3 and C6 appear to vary extensively from session to session; none of these categories exhibit any apparent trends. Because of the fact that we had only one No. 3 operator student whose communications were observed only over three sessions, the results with this subject are quite ambiguous (see Table 9).

Changes in the Direction of Communication. Figure 16 presents the percentage of statements involving student No. 4 operators during training as a function of communication relationships among the student, TACCO/pilot, and instructor (communication between the No. 4 and No. 3 operators did not occur in the sampled sessions). Communications described as "Instructor" include only those statements made to or by the instructor in that capacity. Statements made to or by the instructor as a member of the S-2E crew are described by the pertinent position (e.g., statements made by the instructor in the role of TACCO are described as TACCO communications). The percentages for operational personnel communication exchanges involving the No. 4 operator are presented for the sake of comparison. Two of the four communication exchanges (Figure 16-2, No. 4 operator → TACCO/pilot, and Figure 16-3, Instructor → No. 4 operator) exhibit a trend over sessions toward the communication levels for No. 4 operator personnel. The trend for the operator to increase the proportion of communications to the TACCO function, while a decreasing proportion of the communications are contributed by the instructor, indicates that the communications aspects of the training missions did improve as a function of training; i.e., the communications behavior on the part of the student No. 4 operator began to approach that practiced by operational personnel.

TABLE 9. STATEMENT PERCENTAGE BY CATEGORY FOR THE NO. 3 OPERATOR STUDENT

Sessions	Communication Categories 1	2	3	4	5	6	Σ
4	5	2	42	11	29	11	100%
5	5	5	48	13	30	--	101%
6	--	--	100	--	--	--	100%
Operational No. 3 Operator Average	17	14	43	1	13	11	99%

Figure 17 depicts communications for student TACCO/pilot interactions with the No. 3 operator, No. 4 operator and instructor. Although this figure is based on small Ns (1 and 2) it appears the TACCO/pilot communications to the No. 4 operator (Figure 17-3) may become relatively greater with training, approaching the operational level. (The fact that the percentage for the 7th session is in excess of the operational level could have resulted from the lack of communication with the No. 3 operator position during that session.) It is further apparent that communications to and from the instructor decreased to zero over the training sessions. A note of interest is the comparison of instructor communications percentages given in Figures 16 and 17. The indicated predominance of instructor interaction with the No. 4 operator as compared with the TACCO was corroborated by the observations of the data collectors.

Conclusions Derived from the Communications Analyses. What do the communications analyses tell us about the use of the trainer to improve crew coordination? If one assumes that change in communications pattern is a valid indication of the level of crew coordination, it appears the device does train both the student No. 4 operator and student TACCO/pilot to approximate certain characteristics of the communication behavior as exhibited by operational crews in that same environment. The communications analyses indicate the following changes occurred over the training sessions observed: (1) interaction between the TACCO and No. 4 operator and their use of C4 statements (Verifies Information, Feedback), increased toward the operational team level; (2) interaction with the instructor by both the TACCO/pilot and No. 4 operator decreased; and (3) the relative use of C3, Declarative Information, and C5, Acknowledgments, on the part of the No. 4 operator and TACCO/pilot, respectively, decreased toward the operational team level. It was noted that the instructors tended to communicate more with the student No. 4 operator than with the TACCO; whether or not this is relevant to the erratic performance data for the TACCO position presented earlier should perhaps be considered. Nothing can be said about the No. 3 operator position because of the paucity of information.

THE SEQUENCING OF SIMULATOR SESSIONS WITH FLIGHT TRAINING. Since flight training at Key West for students was interspersed with simulator training sessions, it is reasonable to ask what effect flight training had on trainer performance. One cannot ask this question of San Diego training because in their training regimen, flight training followed completion of the initial training sessions.

This question is of some importance, because to our knowledge there has been little or no research on the effect, if any, of flight training on ground training (the existing transfer of training studies have considered only the reverse case). It is possible to hypothesize that the effect of flight training on performance in the trainer would be to facilitate the

latter, because presumably the same behaviors are being practiced in the air as on the ground. Alternatively, the effect may be inhibitive, if certain of the behaviors practiced in flight were somewhat different from their ground counterparts or important environmental variables were not simulated exactly. If degraded student performance occurs in the simulated environment following a flight session, it would suggest the schedule for trainer versus flight sessions should be evaluated to determine the more efficient and effective schedule.

Unfortunately the investigators cannot provide a definite answer in this study due to the paucity of data. A study of the problem requires: (1) repetitive measures on a sufficient number of students who have had none, one, or more flights intervening between successive trainer sessions with measures in both environments, and (2) the control of trainer sessions so that affective variables extraneous to the question do not confound the results. The inconsistent manner in which flights were scheduled, the small population size, and the incomplete data sets due to incomplete missions make it impossible to provide any data except for the No. 4 operator on Measure C (Figure 18). Although the sample is extremely small in size, it is of particular interest due to the relative stability of other variables (no team sessions, similar lessons, and training schedules).

Although the number of cases (N = 4) is far too small to be conclusive, Figure 18 suggests that the effect of intervening flight training is to reduce operator performance in the trainer on the following session. When student performance without intervening flights (students W and Y) is contrasted with student performance with intervening flights (students T and X), there appears to be either some loss or less improvement in subsequent trainer performance as a function of the intervening flight (the letters T, W, X, and Y identify students in Table 4).

These very tentative results suggest that one should consider not only the consequences of transfer from the trainer to an operational training situation, but also the reverse case: the consequences of transfer from the flight environment to ground training. It may very well be that the resulting efficiency and effectiveness of a training program for performance in an operational environment is, at least in part, an interactive function of the iterated effects from both types of training transfer.

It should be noted here that Measure C for the No. 4 operator was presented in Section II as a demonstration of the effectiveness of continued trainer use to improve performance. Intervening flights were found to affect performance levels in the trainer but, as reference to Figure 18 indicates, the major positive affector of performance was continued trainer use.

HOW DO INSTRUCTORS EVALUATE TRAINER EFFECTIVENESS?

At the conclusion of the study, instructors at San Diego and Key West were asked to evaluate the effectiveness of the training device (see Appendix B). Two types of questions were asked: quantitative and qualitative.

QUANTITATIVE RESPONSES. Questionnaire items requiring a quantitative answer are summarized below:

a. Rate the effectiveness of the position or station in the trainer (Question 1).

b. Rate the overall effectiveness of the trainer (Question 2).

c. Rate the trainer in terms of its effectiveness of crew training (Question 7).

d. Rate the effectiveness of the mission problems presented and their characteristics (Question 8).

e. Rate the realism with which controls and displays are simulated (Question 9).

f. Rate the realism of the information displayed by the trainer (Question 10).

g. Rate the realism of the control forces and the manner in which the trainer simulates the S-2E aircraft (Question 11).

h. Compare the effectiveness of the trainer as contrasted with an equivalent amount of flight training (Question 13).

The scale used for these ratings was a 7-point scale, "1" being low and "7" high, with "4" the midpoint between low and high effectiveness.

Table 10 presents the mean values given by instructors at San Diego to each question in terms of the crew position for which the instructor was evaluating the trainer. The range of responses made is also indicated, as are values for each question and for each crew position. Instructors rated only those positions which they instructed, i.e., No. 3 or No. 4 operator, or TACCO/pilot. The reason is that enlisted instructors train only No. 3 and No. 4 operators; since they are not officers, they have no pilot experience and hence cannot instruct TACCOs and pilots. Officer instructors train the latter. Note that for certain questions not all instructors responded. In addition, in one or two cases the same instructor rated the trainer for two crew positions. Table 11 presents the responses supplied by the Key West instructors.

TABLE 10. INSTRUCTOR EVALUATION OF TRAINER—SAN DIEGO

Questionnaire Items	Pilot	TACCO	No. 3 Operator	No. 4 Operator	Mean Value
1. Rated effectiveness of the position or station in the trainer	3.9** (1.6-5.0)	6.9 (6.6-7.0)	6.5	6.3 (6.0-6.5)	5.9
2. Rated overall effectiveness of the trainer	3.6 (1.5-5.3)	6.6 (5.7-7.0)	6.5	6.8 (6.5-7.0)	5.9
7. Rated effectiveness of trainer for crew training	7.0* (7.0-7.0)	6.7 (6.0-7.0)	6.5	6.8 (6.5-7.0)	6.8*
8. Rated effectiveness of the mission problems and their characteristics	6.8* (6.6-7.0)	6.7 (6.4-7.0)	6.5	5.8 (5.0-6.5)	6.5*
9. Rated realism with which the controls and displays are simulated	4.6 (2.3-6.0)	6.7 (6.2-7.0)	6.5	6.3 (6.0-6.5)	6.0
10. Rated realism of the information displayed by the trainer	5.6* (5.0-6.1)	6.5 (6.1-7.0)	6.5	6.3 (6.0-6.5)	6.2*
11. Rated realism of the control forces and manner in which S-2E aircraft is simulated	1.9 (0.7-3.0)	6.1 (5.2-7.0)	2.5	--	3.5
13. Rated effectiveness of the trainer as contrasted with an equivalent amount of flight training	3.1* (3.0-3.2)	6.0 (5.0-6.8)	5.5	6.3 (5.5-7.0)	5.2*
Mean Value	4.6	6.5	5.9	6.4	
Number of instructors responding	N = 3	N = 3	N = 1	N = 2	

*Not all instructors respond l

**Legend: Mean range (possible range: 1 = low to 7 = high)

TABLE 11. INSTRUCTOR EVALUATION OF TRAINER—KEY WEST

Questionnaire Items	Pilot	TACCO	No. 3 Operator	No. 4 Operator	Mean Value
1. Rated effectiveness of the position or station in the trainer	4.5* (4.0-5.0)	5.5	3.7 (2.0-5.0)	2.8 (1.5-4.0)	4.1
2. Rated overall effectiveness of the trainer	4.4 (4.0-4.8)	4.5	3.2 (2.0-4.2)	3.2 (2.5-4.0)	3.8
7. Rated effectiveness of trainer for crew training	3.4 (1.7-5.0)	5.5	3.1 (2.0-4.0)	3.9 (3.3-4.5)	4.0
8. Rated effectiveness of the mission problems and their characteristics	3.3 (1.6-5.0)	5.5	1.6 (1.2-1.9)	2.2 (1.7-3.0)	3.2
9. Rated realism with which the controls and displays are simulated	4.3 (4.0-4.5)	5.5	4.6 (3.0-5.8)	4.6 (3.2-7.0)	4.8
10. Rated realism of the information displayed by the trainer	5.0 (4.0-6.0)	5.5	3.2 (3.0-3.3)	2.8 (2.2-3.8)	4.1
11. Rated realism of the control forces and manner in which the S-2E aircraft is simulated	1.7 (1.0-2.4)	1.5	2.5 (1.0-3.4)	--	1.9
13. Rated effectiveness of the trainer as contrasted with an equivalent amount of flight training	3.4 (2.0-4.8)	5.5	2.7 (1.5-3.6)	3.2 (2.5-3.5)	4.9
Mean Value	3.8	4.9	3.1	3.2	
Number of instructors responding	N=2	N=1	N=3	N=3	

*Legend: Mean range (possible range: 1 = low to 7 = high)

With few exceptions, the ratings of trainer effectiveness for San Diego instructors are quite high, indicating a very positive attitude toward the device. One major exception is the pilot's position, which is rated substantially below the other positions. The reason for this appears to be the relative lack of fidelity with which the pilot's position is simulated in the trainer; the device does not fly like a plane. On the other hand, there is really no reason why it should, since it was developed as a WST for detection, tracking, and attack, not as a flight vehicle. As a consequence, the pilot position in the trainer is rated (Question 13) as having less effectiveness than an equivalent amount of flight training.

The other positions in the simulator are considered to be simulated quite well at San Diego (see Questions 9 and 11 particularly). The only exception here is the low rating for realism of control forces, etc., given to the No. 3 operator position.

When we come to the Key West evaluation, the situation changes radically. The device as a whole is evaluated as being much less effective by Key West instructors than it was by San Diego instructors. Further, they tend to evaluate trainer effectiveness for the positions differently, the only point of agreement being the relatively high ratings given for the TACCO position. There may be a multiplicity of reasons why this should be so, but one possibility became apparent upon comparison of the qualitative responses made by the instructors at the two locations. This is discussed below.

QUALITATIVE RESPONSES. Instructors were encouraged to list those characteristics of the trainer device which they considered most and least effective. A summary of San Diego and Key West instructor comments per crew member position is given in Table 12. Since the comments made are self-evident, they will not be repeated here. It is worthwhile, however, to note that to the instructors the device does provide an opportunity to practice essential operations in a cost-effective environment.

There were two general areas of complaint: (1) incomplete realism in the simulation of controls, displays, and stimuli presented to the crew member and (2) the frequency of equipment malfunctions which either impaired the instructional function, interrupted the training session, or degraded the mission presentation to the student. With regard to realistic simulation, several specific recommendations were made by the instructors for improvements (primarily equipment modifications and additions) which they felt would enhance trainer effectiveness. With regard to equipment malfunctions, examples respective to the effects noted above would be: the failure to operate or a nonsystematic error in the repeater equipment at the instructor station (making evaluation of student performance difficult); "computer dumping," a term used to

TABLE 12. TRAINER EFFECTIVENESS CHARACTERISTICS

Pilot Position

Effective Characteristics

1. Instrument scan behavior can be developed
2. Emergency and checkout procedures can be taught

Ineffective Characteristics

1. Flight simulation not realistic
2. Lighting simulates night rather than day flights

TACCO Position

Effective Characteristics

1. Close resemblance to aircraft environment for TACCO position
2. Ability to practice ASW tactics

Ineffective Characteristics

1. Equipment unreliability and computer dumping
2. Poor pilot simulation distorts TACCO performance

No. 3 Operator

Effective Characteristics

1. Equipment and layout closely simulate aircraft environment
2. Acquaints student with ASW operations

Ineffective Characteristics

1. MAD situation unreal (displays and control forces)
2. Radar targets unrealistic
3. Computer errors in ECM
4. Cannot monitor student's performance adequately

No. 4 Operator

Effective Characteristics

1. Flexibility and versatility of device

TABLE 12. TRAINER EFFECTIVENESS CHARACTERISTICS (continued)

Ineffective Characteristics
1. Signals unrealistic (e.g., too clean)
2. GRAM selection for LOFAR/CODAR is too limited
3. Insufficient central recorder information for instructor evaluation
4. Computer dumping (i.e., problem erroneously reset to zero)

TABLE 13. FREQUENCY ANALYSIS OF THE AFFECTIVE COMMENTS (E.G., EFFECTIVE, USEFUL, GOOD) MADE BY SAN DIEGO AND KEY WEST INSTRUCTOR PERSONNEL IN THE TRAINER EFFECTIVENESS EVALUATION QUESTIONNAIRE

	San Diego	Key West
Positive	38	41
Negative	31	51
Total	69	92

TABLE 13-1. RAW FREQUENCY DATA

	San Diego	Key West
Positive	55	45
Negative	45	55
Total	100%	100%

TABLE 13-2. PERCENTAGES, POSITIVE VERSUS NEGATIVE

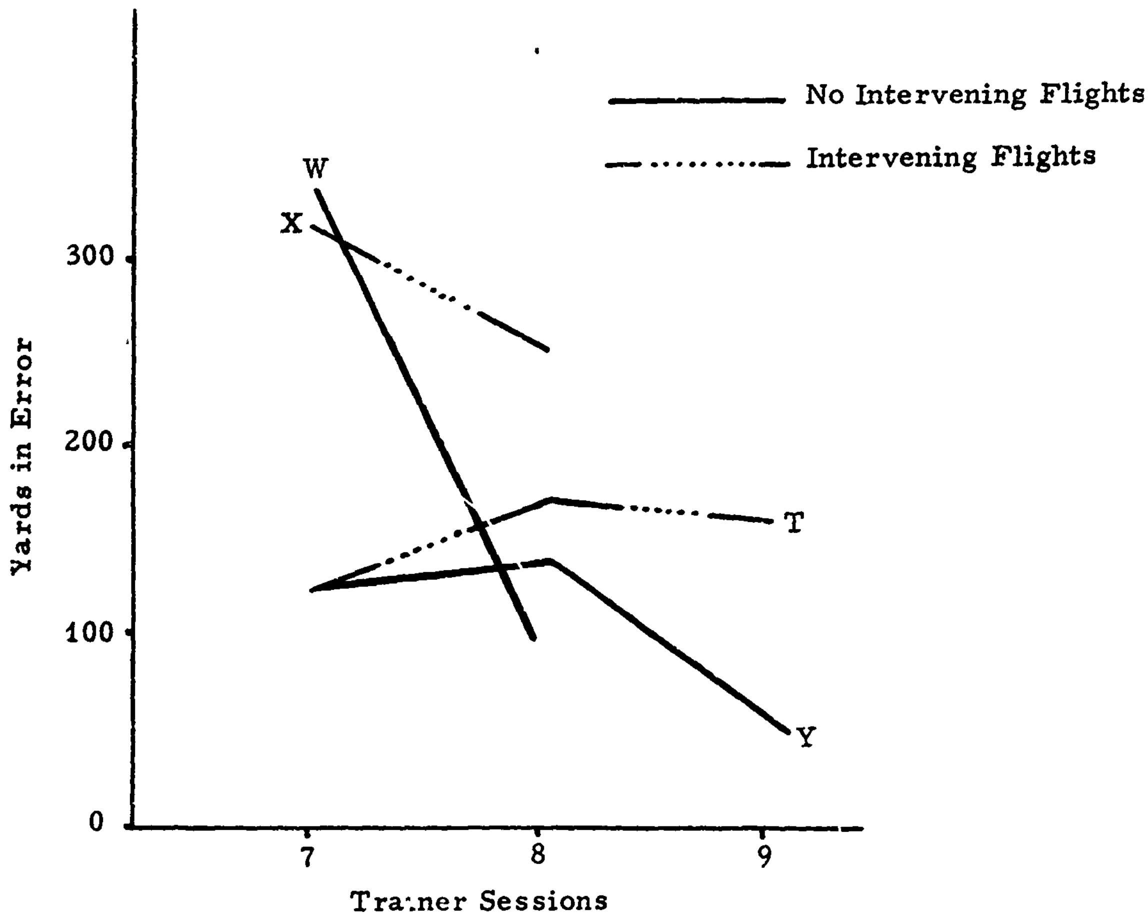


Figure 18. The Effect of Intervening Flights on Student No. 4 Operator Performance, Measure C: Range Error to Target During JULIE (Letters Identify Students Listed in Table 4)

describe computer return of a mission problem to its starting point prior to problem completion; and inadequate or nonperformance of a portion of the student control-display interface, such as drifting of the signal on the radar scope or failure of the student's recording equipment. Occasionally the result of such malfunctions was termination of a training exercise; usually the result was the continuation of the exercise under conditions that were more or less degraded from an instructional viewpoint.

An overview of the affective reaction of the instructors to questions intended to identify the effective and ineffective characteristics of the trainer is provided by a frequency analysis of positive and negative responses made by San Diego and Key West instructors. (Only comments using affective terms, e.g., good or bad, were counted; specific recommendations with no affective overtones were not included.) Both the raw frequency data and the representative percentages are presented in Table 13. Again the more positive reaction of San Diego personnel is evident. It became apparent from a review and comparison of the specific comments made at the two locations that the effects of equipment malfunctions were far more severely felt at the Key West location, where repeated comments were made regarding inadequate maintenance and supply of spare device parts. It is possible that the trainer would be evaluated more highly and positively in Key West given improved equipment operation such that the device capabilities could be utilized more fully.

SECTION III

CONCLUSIONS

The performance measurement data obtained in the course of this study indicate that Device 2F66A is effective for training No. 4 operator personnel in specific duties and tasks; i.e., consistent performance improvements were found to result from continued trainer use. Limited evidence indicates performance on No. 3 operator tasks tends to improve by training on Device 2F66A. The available performance data for the TACCO position, on the other hand, indicates that stable performance levels do not result from the present trainer program, particularly for student personnel (pp 37-52).

The performance data indicated that students and operational personnel benefit from training on Device 2F66A to a larger degree than do reservists. The extreme range in operational personnel performance levels and the indicated improvements in operational crew performance as a result of continued trainer use supports the need for more systematic use by operational as well as student personnel (p 53).

Analysis of communications data indicates that the trainer is effective for training both the student TACCO and student No. 4 operator to modify aspects of their communications patterns in the direction of those exhibited by operational teams in the same setting (pp 60-68).

Syllabus content notwithstanding, training sessions do not systematically increase in complexity and difficulty as defined by training inputs made by the instructor. For the student population, the instructor's desire to individualize training has resulted in large, erratic variations in difficulty; this occurs both between sessions and among students in the same session. For the operational and reserve populations, training device personnel and instructors have little or no control over the type, complexity, or difficulty level of problems used; this combined with irregular device utilization by operational and reserve personnel precludes systematic training for these populations (pp 53-57).

When training No. 4 operator students, the device is used primarily for individual training and only secondarily for full crew training. The presence or absence of full crew training for the TACCO student varies as a function of location. Operational and reserve personnel, however, train as crews (p 58).

The first training session in which full team training is introduced to the No. 4 operator student has a tendency to degrade student performance on the first task performed. It appears that other student task performance observed during the first half of this training mission were affected, but that the nature of the response (improved versus degraded performance) was highly idiosyncratic. Effects on task performance during the latter half of the first team training session were not observable (pp 58-60).

There is a possibility that interweaving flight training sessions has an inhibitive effect on performance in the trainer. Since the utility of the training program to the operational environment may be an interactive function of transfer from both the trainer to flight and vice versa, it is suggested that the possibility of inhibitive effect be further investigated (pp 68-69).

Although all instructors consider the trainer to be useful and necessary, instructors at San Diego rate the trainer more effective than do Key West instructors. Both groups tend to rate the trainer most effective for the TACCO position, and differ in rating trainer effectiveness for the other positions. The majority of criticisms relate to equipment malfunctions and to the insufficient fidelity of some trainer inputs. It appears that the lower ratings given by the Key West instructors may have resulted at least in part from a higher frequency of equipment malfunctions which interfered with instructional goals (pp 70-78).

SECTION IV

RECOMMENDATIONS

In the body of the report, it was suggested that the manner in which Device 2F66A is utilized and the conditions under which it is operated tends to reduce the amount of training it could provide. As a consequence, the trainer fails to achieve its full potential. It is of course only a hypothesis that trainer effectiveness can be improved by modifying trainer use conditions, until learning under different conditions is measured and compared with the results of the present study.

In this section we present certain suggestions for improving trainer use. The recommendations are presented under four headings: (1) Device Use for Student Training, (2) Device Use by Operational Personnel, (3) Trainer Capabilities, and (4) Research Requirements and Evaluation Methodology.

Most of the recommendations for improved use of the device for student training are directed toward modification of the present syllabus. It would be presumptuous of the authors to present these suggestions as a revised syllabus; detailed syllabus revisions should be made only by subject-matter specialists, i.e., S-2E operational and instructional personnel. Hence the suggestions are only presented in general form.

DEVICE USE FOR STUDENT TRAINING

TRAINING INPUTS. The training inputs described in this report should be incorporated into the training schedule in a more systematic manner to create progressively more difficult problems for the student. While the present syllabus does specify the use of some training inputs (e.g., equipment failures) in the training program, the way in which they are actually incorporated was found to be highly variable. Training inputs can serve at least two purposes: (1) realistic practice on reasonable but increasingly difficult mission problems, and (2) development of the task and crew coordination skills necessary to deal effectively with contingencies. It is suggested that training inputs should be incorporated into the student syllabus to effectively serve these two purposes. For example, it is recommended that earlier student training sessions involve a smaller number of training inputs, and that later sessions include combinations of variables, particularly emphasizing multiple targets, target course changes, and target depth changes. The exact combinations should be developed by S-2E operational/instructor personnel, following adaptive training principles and guided by well defined training objectives. Utilizing combinations of variables in a more systematic manner should enhance: (1) the positive transfer of training effects, and (2) the probability that students will receive more complete training, i.e., preparation for a wider range of operational situations.

FULL CREW TRAINING SESSIONS. Assuming that a large enough pool of students is available and that scheduling arrangements can be made, more "full crew" training sessions should be incorporated into the syllabus. Although the practice of using instructors to simulate other crew positions is acceptable and, in fact, desirable in both earlier and supplementary sessions, it cannot be considered completely satisfactory in later sessions that are a part of the basic training program. The full crew training session, it must be remembered, represents the actual operational environment. If remedial training on "part" (individual) aspects of the mission is required, it should be added onto the syllabus and performed at sessions other than those required for performing the basic syllabus.

FURTHER TACCO TRAINING. The results of the present study suggest that student and many of the operational TACCOs require training beyond that observed, and/or training of a different nature. Possible modifications to the TACCO training program might be: (1) more consistent use of the same instructor for a given student rather than the frequent use of different instructors, (2) supplemental part task training on a regular basis, and (3) supplemental full mission lessons presented in individual training sessions, i.e., where every position including that of pilot is enacted by instructor and support personnel. Recommendations (2) and (3) are predicated on this assumption: that demands on the TACCO in his role of team decision-maker and leader may be sufficiently great during periods of mission stress to adversely affect performance of routinely required task behaviors when not sufficiently practiced.

It is entirely possible that the TACCOs are presently learning to perform with respect to criteria not measured in this study. However, if the criteria selected for this study are meaningful operationally, then more effective use of the trainer for the TACCO position may imply additional trainer use, possibly in a somewhat different manner.

MISSION PHASE COVERAGE. Although the present syllabus does its best to cover all mission phases involved in performance of the S-2E mission, not all trainees may receive the same mission coverage. Mission phases may be selected for training on the basis of perceived student needs, and repetitive practice may be given on those phases in which the student appears deficient. As pointed out earlier, remedial training should, if possible, add to the training program rather than substitute for certain mission phases. Students should have an opportunity to practice all mission phases adequately. Review of the comparative emphasis on each of the mission phases during operational performance evaluations, and consideration of the phases more difficult for the student to learn, may indicate the relative amount of practice advisable for each mission phase. It is suggested that additional efforts be made to bring more of the crews through the terminal or "kill" (MAD search and attack phases) part of the mission. There are two reasons for this: (1) both the crew and the instructor then have one objective index as to how well

the members have performed as a team (i.e., in terms of kill and miss distance), and (2) it was observed that the possibility of a kill definitely provided a motivating goal for S-2E personnel. This latter suggestion is of course tied in with the inclusion of more full crew sessions in the syllabus.

DEVICE USE BY OPERATIONAL PERSONNEL

Recognizing that the training of operational personnel is largely controlled by the operational squadrons, efforts should be made to schedule these personnel more systematically, so that all teams receive more frequent sessions in the trainer. It would, however, be pointless to suggest standardized refresher training for such personnel unless arrangements could be made organizationally to subject them to a standard syllabus. There is no question in our minds, on the basis of collected data, that operational personnel would benefit from such standardization and from more frequent sessions.

TRAINER CAPABILITIES

TRAINER MAINTENANCE. The present level of trainer maintenance and unavailability of spare parts was found to interfere with the training program, making systematic training difficult to achieve. It also tended to degrade acceptance of the training device on the part of all personnel, especially at the Key West training site. It is strongly recommended that steps be taken to reduce the current prevalence of equipment malfunctions (e.g., computer dumping, repeater errors, scope irregularities) which interfere with the instructional function and reduce acceptance of the device.

TRAINER FIDELITY. Although there are limits to the extent that trainer fidelity can be cost-effective, it may well be that the fidelity of certain characteristics of Device 2F66A could be reasonably enhanced. The resulting increase in trainer effectiveness would probably be worthwhile. It is suggested that requirements and costs be investigated for increasing the fidelity of the signal presentations made to the No. 3 and No. 4 operators. Also is reasonable, signal presentations should be upgraded to more closely approximate signals found in the operational environment. Further, the training input option of multiple targets, not presently available at Key West, should be provided there.

HARDWARE CAPABILITY. Although it is understood that the S-2E aircraft will be gradually phased out of operation, it is expected that they will be used for some time to come; if this is the case, then S-2E personnel should be trained to use the equipment capabilities available to them on board the aircraft. To do this, the hardware capability of the present S-2E trainers should be updated to permit realistic practice with,

for example, active as well as passive sonobuoys. It is suggested that the objections to the trainer voiced by instructors and by operational personnel be closely evaluated to determine reasonable modifications which would effectively increase the hardware capabilities of the training device.

RESEARCH REQUIREMENTS AND EVALUATION METHODOLOGY

The first four recommendations given below describe areas where further research is needed. Three of the four recommendations (items 1, 3, and 4) are necessary not only for the development of a more effective S-2E training program, but also for the evaluation of the validity of certain basic concepts on which many military training programs are based (but which have not been demonstrated empirically as appropriate or adequate to the task in their present form). In all four cases the type of research program required to provide definitive answers is a comparative evaluation of student performance under alternative training programs and under operational flight conditions. It is only in this manner that the more effective training program for operational requirements can be defined, and useful guidelines for the design of future training programs established.

The fifth recommendation discusses requirements for an effective evaluation methodology based on what has bee learned from this study.

RECOMMENDATION NO. 1. The training inputs utilized by S-2E instructor personnel can be, and in other contexts, usually are considered adaptive training variables. Although conclusive results could not be derived from the available data, it appears that the effect of these training inputs is highly idiosyncratic. That is, the extent and the nature of their effect appears to be a complex function of individual student characteristics, point in training, concurrent events, and the task implications of the input at a particular point in the mission.

Further, interrogation of the individual instructors and data collection personnel concerning the impact of the training inputs on student performance parameters made it evident that evaluation of the impact was almost as idiosyncratic as the impact itself. Under these circumstances it would appear that the development or implementation of a training program according to the adaptive training principle of increasing difficulty on the basis of judgment alone cannot be expected to result in an optimum program. Guidelines based on empirical evidence are required if the potential of adaptive training variables are to be realized. It does appear entirely possible to conduct empirical investigations of training input effects. If the present study, for example, had had a larger student population and control of variables such as equipment malfunctions and the scheduling of the training inputs, it would have provided a considerable amount of information on this question. It is

suggested, therefore, that future investigations based on reasonably complex training devices (such as the S-2E trainer) be conducted for two purposes: (1) definition of those parameters that may affect the action of adaptive training variables on student performance (e.g., individual student characteristics, point in training), and (2) definition of the actual action of adaptive training variables on student performance (e.g., it would appear that the effect on performance may be facilitative if presented at the appropriate time). It is reasonable to expect adaptive training variables to possess tremendous potential for increasing the effectiveness of training programs, but the present study suggests that this potential will not be fully realized unless the variables are more systematically incorporated into these programs.

RECOMMENDATION NO. 2. Modifications of the student TACCO training program were suggested earlier. Considering the several changes that could be made, and the insufficient amount of data provided by the present study, it is recommended that comparative evaluations be made for any new TACCO programs that are developed.

RECOMMENDATION NO. 3. It was previously suggested that student No. 4 operators would benefit from additional team training sessions, while, on the other hand, more individual training was suggested for student TACCO training. Actually it may be constructive to question the concept that team training is necessarily useful, and to question when it is most effective. The current crew training facilities might be more fully utilized and the training programs more effective if the comparative utility of individual versus team training for alternative points in the training program and for various crew member tasks and functions were better understood.

RECOMMENDATION NO. 4. The evidence from the present study suggests that the practice of introducing the student to the operational environment prior to completion of ground training should be questioned and the effects on performance in the operational environment investigated. The effectiveness of training programs for operations in the flight environment may very well be a function of flight versus trainer lesson sequencing. If this is the case, investigations into the nature of the relationship could be well worth the effort.

RECOMMENDATION NO. 5. The advantages and the disadvantages of an evaluation conducted on a noninterference basis were discussed in an earlier section: the questions of realistic and valid data, contrasted with the difficulties of analyzing such data so as to obtain valid answers. It can be noted that if the present study had been run in a completely controlled manner and the experimental design (i.e., alternative conditions) had not been quite so complex, it is entirely possible that the research questions posed by items 1, 3, and 4 above might never been

raised. Further, a realistic picture of the training provided by the program as it is presently operated would not have been provided. These are substantial and worthwhile gains obtained from this study. Basic concepts have been opened to question while the inadequacy of a program in its present form has been demonstrated.

On the other hand, it is entirely evident to the authors that, given certain conditions, a better and more complete evaluation of performance changes resulting from continued trainer use would have resulted. The primary condition is the requirement for a larger population sample with a constant membership for all of the sessions. For example, the evaluation of both TACCO and No. 3 operator performance in this study was extremely constricted by both the small student population and the incomplete coverage of the possible student sessions (i.e., performance was not measured in either the earliest sessions or the most advanced sessions). Further, the more variables one wishes to analyze for possible concurrent and/or interactive effects on personnel performance, the larger the population size required for analysis purposes.

A second condition which can be helpful is the opportunity for repeated measurement of student performance on one or more standardized lessons. The availability of such a measurement opportunity is usually dependent on such things as number of lessons in the program, lesson length, scheduling problems, and whether or not any investigator interference with the training program is permitted. It should be noted, however, that such data is not sufficient in and of itself. A study which presents data collected only on a standardized lesson, when this is not normal training procedure and/or the student is receiving considerable training in other lessons which are not standardized, does not really demonstrate anything about the effectiveness of the device as it is normally used. Such data, however, presented in conjunction with data collected on a noninterference basis, does serve to corroborate the tendencies toward improved performance displayed by the noninterference data curves.

REFERENCES

Briggs, G. E. and Johnston, W. A. Team Training. NAVTRADEVCEN 1327-4, June 1967, Naval Training Device Center, Orlando, Florida.

Federman, P. and Siegel, A. I. Communication as a Measurable Index of Team Behavior. NAVTRADEVCEN 1537-1, October 1965, Naval Training Device Center, Port Washington, New York.

Jeantheau, G. Training System Use and Effectiveness Evaluation. July 10, 1970, Naval Training Device Center, Orlando, Florida.

Krumm, R. L. and Buffardi, L. Training Effectiveness Evaluation of Naval Training Devices. Part I: A Study of Submarine Diving Trainer Effectiveness. NAVTRADEVCEN 69-C-0322-1, May 1970, Naval Training Device Center, Orlando, Florida.

Krumm, R. L. and Farina, A. J. Effectiveness of Integrated Flight Simulator Training in Promoting B-52 Crew Coordination. MRL-TDR-62-1, February 1962, Behavioral Sciences Laboratory, 6570th Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

Metersky, M. L. A Study of Simulated ASW Crew Performance in Relation to Selected Fatigue Factors. NADC-SY-6709, December 1967, Naval Air Development Center, Johnsville, Pennsylvania.

APPENDIX A

NTDC STUDY

TRAINER EFFECTIVENESS

EVALUATION

S-2E TRAINER

CREW PERFORMANCE

RATING PACKAGE

Mission No. ____________ Pilot ____________ #3 ____________

Date ________________ TACCO ____________ #4 ____________

Instructor ________________________________

I. OVERALL CREW PERFORMANCE

Using the scale presented below please rate the overall crew performance during this training exercise. You may place an (x) at any point along the continuum.

Minimal Proficiency — Above Average

II. OVERALL INDIVIDUAL CREW MEMBER PERFORMANCE

For each of the crew positions (TACCO, #3 Operator and #4 Operator) please rate each individual crewmember using one of the three categories given. After you have selected one of the categories, then rate the individual as to where you feel his performance falls within that category using the scale presented.

A. TACCO

Please rate the overall performance of the TACCO based on his selection of the appropriate strategy, laying of the proper buoy pattern, operation of the ASN/30, etc.

a. AA _____
b. A _____
c. BA _____

Low — High

B. #3 Operator

Please rate the overall performance of the #3 Operator based on his speed and accuracy of detection and classification, reporting speed, and the amount of update information provided, etc.

a. AA _____
b. A _____
c. BA _____

Low — High

C. #4 Operator

Please rate the overall performance of the #4 Operator based on his speed and accuracy of detection and classification, reporting speed, and the amount of update information provided.

a. AA ____
b. A ____
c. BA ____

Low High

III. PERFORMANCE IMPROVEMENT

A. Overall Crew Improvement *

For the crew (assuming that the three relevant positions have trained previously as a unit, and excluding the pilot) indicate, using the scale below, the proficiency of the crew during the previous exercise in which they performed as a unit (use an "o") and their proficiency during this exercise (use an "x").

Low High

B. Individual Crewmember Improvement

For each of the three crew positions, perform the same procedure as in "A" above.

TACCO Low High

#3 OPR Low High

#4 OPR Low High

* Note If total crew (TACCO, #3 and #4) are not involved, please make sure item B, below is completed.

APPENDIX B

NTDC STUDY
TRAINER EFFECTIVENESS
EVALUATION

S-2E TRAINER

DEVICE RATING PACKAGE

THE BUNKER-RAMO CORPORATION

DEFENSE SYSTEMS DIVISION

31717
LA TIENDA
DRIVE
WESTLAKE VILLAGE
CALIFORNIA 91360
PHONE
213 889-2211

LOCATION		POSITION UNDER EVALUATION	
San Diego (VS-41)	________	_____	Pilot
Key West (VS-30)	________	_____	TACCO
If the location is Key West:		_____	#3
S2E Trainer	________	_____	#4
ASN 30 Trainer	________		

INSTRUCTIONS

As part of an evaluation of training device effectiveness being conducted for the U. S. Navy we would like you to answer the following questions as completely and honestly as you can. This is not a test: there are no right or wrong answers. Rather what we are after is your carefully considered responses based on your own experience and competence.

Some of the questions require two kinds of answers; a rating of the device or its characteristics and your own explanation of that answer. Please be as detailed as you can.

TRAINEE'S STATION EVALUATION

1. Rate the effectiveness (that is- how much does it contribute to learning and training) of the trainee's station?

Low High

Explain here if you wish:

2. Overall, how effective would you say the training device is as a trainer for this position?

Low High

3. What characteristics of the device would you say contribute to its effectiveness or ineffectiveness?

Ineffectiveness	Effective

4. In terms of working with the trainee during a training session which aspects of the trainer's design, do you feel, aid or hinder your efforts as an instructor?

Aid	Hinder

5. What changes in the design or operation of the device do you think should be made in order to enhance the training at this station?

6. Are there any modifications or improvements you would suggest that would allow better or more efficient evaluation of student performance at this station (i.e. either in the form of recording devices, better or more slave equipment, etc.)?

7. How effective is the trainer in providing the trainee at this station with the required <u>crew training and coordination</u> that he requires as a member of an operational crew?

Low ——————————————————— High

8. How effective are the problems and the problem characteristics (targets, sea-state, speed changes, etc.) in providing training for this position? Are they best problems for training or could they be improved?

Low ——————————————————— High

9. How realistically are the operating controls/displays simulated for this position?

Low Realism ——————————————————— High Realism

10. Are indicator values and/or displayed information presented realistically and appropriately? What would you suggest as improvements?

Poor Presentation						Good Presentation

Improvements:

11. How accurately are control forces and other indicators of aircraft performance presented to the student at this station?

Low Accuracy						High Accuracy

12. Do you think that changing the situation, described by your answer to question 11, would have any effect on training performance?

13. How effective do you think the trainer is at teaching compared to an equivalent amount of flight training?

Question 14

In this question we would like you to rate the effect of several trainer variables on performance during different mission phases. We realize that each of these variables has a different amount of effect at each position and during each phase. You are asked to indicate those variables that have an effect during a particular task and to evaluate the comparative amount of impact.*

* Indicate the amount of impact the variable has by using the terms Low, Medium, or High. Leave the space blank if there is no relationship.

(Fill out only the matrix appropriate to the position you are evaluating.)

TACCO

Question 14: Evaluation of Variable Impact on Performance

	LOFAR		CODAR		JULIE
Trainer Variables	Time to Detection	Time to Classification	Time to Detection	Procedural Errors	Calibration Range
Multiple Targets (subs)					
Target Alters Speed (rpm)					
Target Alters Course					
Target Alters Depth					
Recorder Failure Induced					
Sonobuoy Failures					
Receiver Malfunction Induced					
Loss of Contact Induced					
Other Crew Members Performance (indicate which member)					

#3

Question 14: Evaluation of Variable Impact on Performance

Trainer Variables	LOFAR/CODAR: Time to Detect Sub	LOFAR/CODAR: ECM/RADAR Contact Classification	LOFAR/CODAR: Contact Range	LOFAR/CODAR: Time Between Updates	MAD: Target Detection
Multiple Targets (subs)					
Target Alters Speed (rpm)					
Target Alters Course					
Target Alters Depth					
Recorder Failure Induced					
Sonobuoy Failures					
Receiver Malfunction Induced					
Loss of Contact Induced					
Other Crew Members Performance (indicate which member)					

#4

Question 14: Evaluation of Variable Impact on Performance

	LOFAR	CODAR			JULIE	MAD ATTACK
Trainer Variables	Sequential Drops: Range, Bearing	Sub Position	Drop Points	Number of CODAR Plants	Pattern Accuracy: Range, Bearing	Kill, No Kill Miss Distance
Multiple Targets (subs)						
Target Alters Speed (rpm)						
Target Alters Course						
Target Alters Depth						
Recorder Failure Induced						
Sonobuoy Failures						
Receiver Malfunction Induced						
Loss of Contact Induced						
Other Crew Members Performance (indicate which member)						